Phylogenetic, biochemical and structural assessment of key enzymes in ectoine and hydroxyectoine biosynthesis

Kumulative Dissertation

zur

Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem Fachbereich Biologie

der Philipps-Universität Marburg vorgelegt von

Nils Widderich

aus Mönchengladbach

Marburg/Lahn, Februar 2017

Phylogenetic, biochemical and structural assessment of key enzymes in ectoine and hydroxyectoine biosynthesis

Kumulative Dissertation

zur

Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem Fachbereich Biologie

der Philipps-Universität Marburg vorgelegt von

Nils Widderich

aus Mönchengladbach

Marburg/Lahn, Februar 2017

Die Untersuchungen zur vorliegenden Arbeit wurden von Oktober 2012 bis September 2015 im Laboratorium für Mikrobiologie am Fachbereich Biologie an der Philipps-Universität Marburg unter der Leitung von Herrn Prof. Dr. Erhard Bremer durchgeführt.

Vom Fachbereich Biologie der Philipps-Universität Marburg als Dissertation angenommen am: 27.04.2017

Erstgutachter: Prof. Dr. Erhard Bremer
Zweitgutachter: Prof. Dr. Andreas Brune

Tag der mündlichen Prüfung am: 11.07.2017

Die während dieser Promotion erzielten Ergebnisse wurden in den folgenden sieben Manuskripten und einem Übersichtsartikel veröffentlicht:

Widderich, N.*, Pittelkow, M.*, Hoeppner, A., Mulnaes, D., Buckel, W., Gohlke, H., Smits, S.H.J. & Bremer, E. (2014). Molecular dynamics simulations and structure-guided mutagenesis provide insight into the architecture of the catalytic core of the ectoine hydroxylase. *J Mol Biol* 426:586-600.

Hoeppner, A., **Widderich, N.**, Bremer, E. & Smits, S.H.J. (2014). Overexpression, crystallization and preliminary X-ray crystallographic analysis of the ectoine hydroxylase from *Sphingopyxis alaskensis*. *Acta Crystallogr F Struct Biol Commun* 70:493-496.

Widderich, N.*, Hoeppner, A.*, Pittelkow, M., Heider, J., Smits, S.H.J. & Bremer, E. (2014). Biochemical properties of ectoine hydroxylases from extremophiles and their wider taxonomic distribution among microorganisms. *PLoS One* 9:e93809.

Hoeppner, A.*, **Widderich, N.***, Lenders, M., Bremer, E. & Smits, S.H.J. (2014). Crystal structure of the ectoine hydroxylase, a snapshot of the active site. *J Biol Chem* 289:29570-29583.

Kobus, S., **Widderich, N.**, Hoeppner, A., Bremer, E. & Smits, S.H.J. (2015). Overproduction, crystallization and preliminary X-ray diffraction data analysis of the ectoine synthase from the cold-adapted marine bacterium *Sphingopyxis alaskensis*. *Acta Crystallogr F Struct Biol Commun* F71:1027-1032.

Widderich, N.*, Czech, L.*, Elling, J.E., Könneke, M., Stöveken, N., Pittelkow, M., Riclea, R., Dickschat, J.S., Heider, J. & Bremer, E. (2016). Strangers in the archaeal world: osmoregulated synthesis of the compatible solutes ectoine and hydroxyectoine by the Thaumarchaeon *Nitrosopumilus maritimus*. *Env Microbiol* 18, 1227–1248.

Widderich, N.*, Rodrigues, C.D.*, Commichau, F.M., Fischer, K.E., Ramirez-Guadiana, F.H., Rudner, D.Z. & Bremer, E. (2016). Salt-sensitivity of SigH and Spo0A prevents sporulation of *Bacillus subtilis* at high osmolarity avoiding death during cellular differentiation. *Mol Microbiol* 100, 108–124.

Übersichtsartikel:

Widderich, N., Bremer, E. & Smits, S.H.J. (2016). The ectoine hydroxylase: a non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase. In: *Encyclopedia of Inorganic and Bioinorganic Chemistry*. (Ed. A. Messerschmidt) Wiley Online Library. doi: 10.1002/9781119951438.eibc2440.

* These authors contributed equally

Meiner Familie

Table of contents

Table of figures

Summary

In their natural habitats, microorganisms have to frequently cope with a multitude of stressful biotic and abiotic conditions that can have adverse effects on their growth and their survival in a given ecosystem. One of the most important parameters of environmental stress due to its impact on almost all microorganisms is changing osmolarity/salinity. Among microorganisms, accumulation of compatible solutes is a widely used strategy to preserve cell integrity and growth under hyperosmotic conditions. The tetrahydropyrimidines ectoine and hydroxyectoine belong to these osmolytes and are produced by many procaryotes to minimize the adverse effects of high osmolarity on cellular physiology. Due to their beneficial impact on macromolecules, ectoine and hydroxyectoine are also reffered to in literature as chemical chaperones. These properties have spurred considerable biotechnological interest in ectoines.

Ectoine is enzymatically synthesized by the ectoine synthase (EctC) and the ectoine hydroxylase (EctD) catalyses the conversion of ectoine to hydroxyectoine. Despite the fact that both enzymes have already been studied somewhat, in-depth knowledge on their phylogenetic distribution, biochemistry and structure was still lacking prior to this dissertation since only a minor number of EctC and EctD proteins has been characterized and crystal structures of both proteins containing all ligands were still missing.

Since a deeper understanding of these key enzymes in ectoine biosynthesis is desirable, both with respect to basic science and industrial applications, the aim of the present dissertation was to assess the phylogenetic affiliation of ectoine biosynthetic genes and to study a selection of EctC and EctD enzymes with respect to their biochemical and kinetic properties. In addition, crystallographic approaches of EctC and EctD and site-directed mutagenesis experiments of EctD were conducted to provide a basis to unravel the position and binding motifs of the ligands within the catalytic cores of EctC and EctD.

To elucidate the phylogenetic distribution of ectoine biosynthesis, the amino acid sequences of both key enzymes were used as a search query and identified, after removal of redundant sequences, about 723 potential ectoine producers of which only 12 originated from *Archaea*. This analysis revealed that ectoine biosynthesis is widely distributed in prokaryotes, predominantly in members of the *Bacteria*, underlining the important role of ectoines in microbial stress responses.

On this basis, various ectoine synthases and ectoine hydroxylases deriving from different organisms have been biochemically and/or kinetically characterized that were

widely distributed on the phylogenetic tree of ectoine biosynthesis. Each of the so far studied proteins possessed similar enzyme kinetics, however, comparison of their biochemical characteristics revealed only minor differences between EctD proteins but major variations between EctC enzymes. This suggests that the ectoine synthase, whose properties partially reflect the environmental circumstances of their hosts, might have developed in terms of evolution prior to the ectoine hydroxylase.

Identification of EctC and EctD proteins possessing increased stability allowed new crystallization trials. Multiple crystal structures have been solved in the course of this dissertation in collaboration with Dr. Sander Smits (University of Düsseldorf). In terms of EctD, structures in its apo-form, in complex with the iron co-factor and in complex with the iron catalyst, the co-substrate 2-oxoglutarate and the reaction product hydroxyectoine have been solved. These structures provided, in connection with comprehensive site-directed mutagenesis experiments, a detailed view into the catalytic core of EctD allowing a proposal for its catalyzed reaction mechanisms (Proposal by Dr. Wolfgang Buckel). In terms of EctC, a detailed expression and purification protocol for the ectoine synthase from the cold-adapted marine bacterium *Sphingopyxis alaskensis* has been described in collaboration with Dr. Sander Smits (University of Düsseldorf) that identifies EctC to form dimers in solution and deals with crystallization trials and preliminary X-ray diffraction data providing a promising basis for the solution of its crystal structure, which has, based on the presented data, been published posterior to the present dissertation.

Collectively, the present dissertation provides detailed information about the phylogenetic distribution and biochemistry of EctC and EctD as well as the solved crystal structure of the ectoine hydroxylase, the key enzymes in hydroxyectoine biosynthesis, and promising crystallization trials of EctC, the key enzyme in ectoine biosynthesis, and thus substantially contributes to the understanding of the role of ectoines in the global microbial osmostress adaptation.

Zusammenfassung

In ihren natürlichen Habitaten müssen Mikroorganismen regelmäßig eine Vielzahl biotischer und abiotischer Stressbedingungen, welche beeinträchtigende Effekte auf ihr Wachstum und ihr Überleben in einem gegebenen Ökosystem haben können, bewältigen. Aufgrund ihres Einflusses auf nahezu alle Mikroorganismen ist wechselnde Osmolarität/Salinität einer der wichtigsten durch die Umwelt bedingten Stressfaktoren. Unter Mikroorganismen ist die Akkumulation von kompatiblen Soluten eine weitverbreitete Anpassungsstrategie um die Zellintegrität und das Wachstum unter hyperosmotischen Bedingungen zu wahren. Die Tetrahydropyrimidine Ectoin und Hydroxyectoin sind solche Osmolyte und werden von vielen Prokaryoten synthetisiert um die beeinträchtigenden Auswirkungen hoher Osmolarität auf die zelluläre Physiologie zu minimieren. Aufgrund ihres positiven Einflusses auf Makromoleküle werden Ectoin und Hydroxyectoin in der Literatur auch als chemische Chaperone bezeichnet. Diese Eigenschaften haben Ectoine bereits für biotechnologische Anwendungsgebiete interessant gemacht.

Ectoin wird enzymatisch durch die Ectoin-Synthase (EctC) synthetisiert und die Ectoin-Hydroxylase (EctD) katalysiert die Umwandlung von Ectoin zu Hydroxyectoin. Obwohl beide Enzyme bereits einigen wenigen Studien unterzogen wurden mangelt es an Detailwissen hinsichtlich ihrer phylogenetischen Verbreitung, Biochemie und Struktur, da nur eine kleine Zahl von EctC- und EctD-Proteinen vor dieser Dissertation charakterisiert wurde und Kristallstrukturen beider Proteine mitsamt aller gebunden Liganden noch immer fehlten.

Da ein vertieftes Verständnis dieser Schlüsselenzyme in der Ectoin-Biosynthese sowohl hinsichtlich Grundlagenforschung als auch mit Blick auf deren industrieller Anwendungsbereiche erstrebenswert ist, war das Ziel der vorliegenden Dissertation die phylogenetische Verbreitung der Ectoin-Biosynthesegene zu taxieren und eine Selektion von EctC- und EctD-Enzymen hinsichtlich ihrer biochemischen und kinetischen Parameter zu studieren. Zusätzlich wurden von beiden Proteinen kristallographische Studien durchgeführt, um die Position und Bindemotive der verschiedenen Liganden in den katalytischen Zentren von EctC und EctD zu entschlüsseln. Diese wurden im Falle von EctD mittels gerichteter Mutagenese-Experimenten unterstützt.

Um die phylogenetische Verbreitung der Ectoin-Biosynthese aufzuklären wurden die Aminosäure-Sequenzen beider Proteine als Suchanfrage genutzt, wodurch, nach Aussortierung redundanter Sequenzen, 723 potentielle Ectoin-Produzenten identifiziert werden konnten, von welchen lediglich 12 aus Archaeen stammten. Diese Analyse

zeigte, dass die Ectoin-Biosynthese unter den Prokaryoten, und dabei überwiegend unter den Bakterien, weit verbreitet ist. Somit konnte die wichtige Rolle der Ectoine in der mikrobiellen Stressantwort unterstrichen werden.

Auf dieser Grundlage wurden verschiedene Ectoin-Synthasen und Ectoin-Hydroxylasen, welche aus unterschiedlichen Organismen stammten und auf dem phylogenetischen Ectoin-Biosynthese Stammbaum weit verbreitet waren, biochemisch und/oder kinetisch charakterisiert. Jedes der studierten Enzyme wies ähnliche Enzymkinetiken auf, der Vergleich ihrer biochemischen Charakteristika offenbarte jedoch geringfügige Unterschiede zwischen den EctD-Proteinen und signifikantere Unterschiede innerhalb der EctC-Enzyme. Dies suggeriert, dass sich die Ectoin-Synthase, deren Eigenschaften partiell die Umweltbedingungen ihrer Wirte reflektiert, evolutionär vor der Ectoin-Hydroxylase entwickelt haben könnte.

Die Identifizierung von EctC- und EctD-Proteinen, welche vergleichsweise erhöhte Stabilität aufwiesen, erlaubte zudem neue Kristallisationsversuche. Mehrere Kristallstrukturen konnten im Rahmen dieser Dissertation und in Kollaboration mit Dr. Sander Smits (Universität Düsseldorf) gelöst werden. Hinsichtlich EctD war es möglich Strukturen des Enzyms in seiner Apo-Form, in Komplex mit dem Eisenkofaktor, und in Komplex mit dem Eisenkatalyst, dem Kosubstrat 2-Oxoglutarat und dem Reaktionsprodukt Hydroxyectoin zu lösen. Diese Strukturen lieferten, in Verbindung mit umfassenden ortsgerichteten Mutagenesestudien, einen tiefen Einblick in das katalytische Zentrum von EctD und erlaubten eine Vorhersage für dessen katalysierten Reaktionsmechanismus zu treffen (Vorhersage von Dr. Wolfgang Buckel). Hinsichtlich EctC wurde ein detailliertes Expressions- und Reinigungsprotokoll von EctC aus *Sphingopyxis alaskensis* in Kollaboration mit Dr. Sander Smits (Universität Düsseldorf) beschrieben, das EctC als Dimer identifiziert und Kristallisationsversuche und vorläufige Röntgenstrukturanalysen-Daten abhandelt, welche eine vielversprechende Grundlage für die Auflösung der EctC Kristallstruktur schaffen, welche aufbauend auf diesen Daten, im Anschluss an die vorliegende Dissertation veröffentlich wurde.

Zusammenfassend liefert die vorliegende Dissertation detaillierte Informationen über die phylogenetische Verbreitung und Biochemie der Schlüsselenzyme der Ectoin- und Hydroxyectoin-Biosynthese sowie die gelöste Kristallstruktur von EctD und vielversprechende Kristallisationsverusche von EctC und leistet somit einen substantiellen Beitrag zum Verständnis der Rolle von Ectoinen in der globalen mikrobiellen Osmostress-Anpassung.

Introduction

I. Changing environments

In their natural habitats, microorganisms are frequently exposed to changing environmental circumstances. Therefore, they have to cope with a multitude of stressful biotic and abiotic factors that can have adverse effects on their growth, proliferation and survival in a given ecosystem. Such inhibitory environmental factors of growth can for instance be variations in the pH value, differences in growth temperature, reactive oxygen species, nutrient limitation or changing osmolarity (Csonka and Epstein, 1996; Kempf and Bremer, 1998; Lucht and Bremer, 1994). Hence, microbial communities thrive in their ecosystems as a consequence of their ability to sense environmental changes and respond to them in a timely and appropriately coordinated manner (Lopez and Kolter, 2010). Most microorganisms have evolved a globally-acting adaptation process to environmental challenges, the general stress response (Storz and Hengge-Aronis, 2000). In addition, many microorganisms have developed stringent stress responses to specific challenges and some are also able to induce several cellular differentiation processes (Storz and Hengge-Aronis, 2000).

II. Cellular differentiation processes

Many bacteria have evolved complex differentiation processes to adapt to changes in their environment. The continuous sensing of the physical, chemical and biological parameters of their surroundings allows microbial populations to respond to environmental changes in a coordinated fashion. Some organisms thereby possess, as a consequence of differential gene expression, the capability to differentiate into subpopulations yielding genetically identical yet phenotypically and/or morphologically distinct cell types. Well studied examples of cellular differentiation processes are fruiting-body formation in *Myxococcus xanthus* (Konovalova *et al*., 2010) and asymmetric division of *Caulobacter cresentus* (Holtzendorff *et al*., 2004).

Probably the best and most-extensively studied differentiating organism is the soil-dwelling bacterium *Bacillus subtilis* (Mandic-Mulec *et al*., 2015). It is an appreciated model organism for cellular differentiation processes since it is capable to display various distinct cell types and is thus often considered as the master of differentiation (Lopez and Kolter, 2010; Lopez *et al*., 2009). In response to different environmental signals, *B. subtilis* is able to differentiate into subpopulations resulting from modified gene expression cascades in these cells (Fig. 1). These subpopulations may consist of flagella-producing motile cells, competent cells able to take up and incorporate free

DNA from their environment, toxin-secreting cells cannibalizing their siblings, extracellular matrix producing cells that form robust biofilms and cells that produce long-lasting and highly resistant endospores (Lopez and Kolter, 2010; Lopez *et al*., 2009). Quite often, the different subpopulations can be distinguished from each other visibly since their altered gene expression results in morphological changes, however, some can only be distinguished by monitoring cell-type-specific gene expression using molecular reporter constructs.

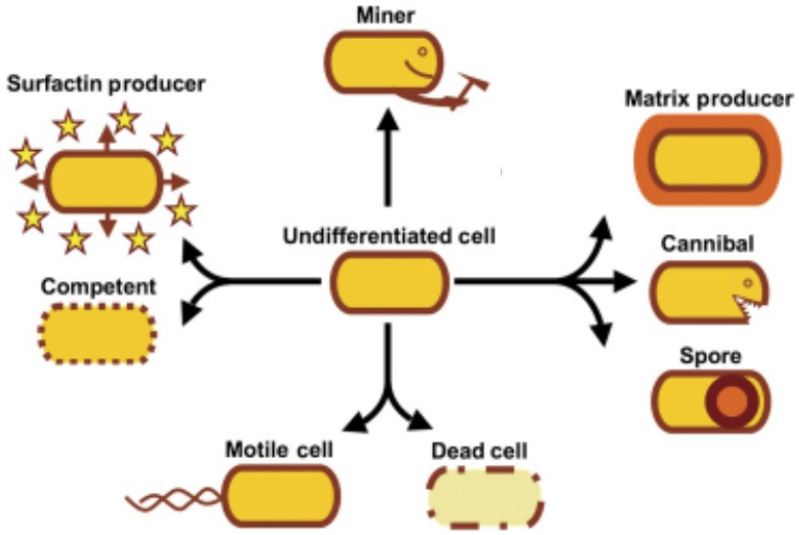


Fig.1. Cellular differentiation processes in *Bacillus subtilis*. Each cell type is schematically illustrated with respect to its most representative attribute. The cell types are divided into subgroups with accordance to the master regulator the corresponding differentiation and black arrows indicate the process of differentiation. Image is adapted and modified from López und Kolter (Lopez and Kolter, 2010).

The probably best and most intensively studied cell fate of *B. subtilis* is spore formation. Sporulation is typically induced in response to nutrient limitation. Thereby, a set of sporulation-inducing kinases initiates a highly orchestrated developmental program that involves approximately a quarter of the genes encoded by the genome of *B. subtilis*. (Eichenberger *et al*., 2004; Errington, 2003; Higgins and Dworkin, 2012; Nicolas *et al*., 2012; Steil *et al*., 2005). During this developmental program, the cell undergoes an asymmetric division that results in the formation of two different cell types, the mother cell and the forespore (Fig. 2). Both follow distinct developmental routes of gene expression driven by stage- and compartment-specific transcription factors which are linked to each other by cell-cell signaling pathways. Finally, upon maturation of the developing forespore, lysis of the mother cell releases a highly stress-resistant endospore into the environment (Errington, 2003; Higgins and Dworkin, 2012; Stragier and Losick, 1996). These endospores can remain dormant in their ecosystems for extended time periods yet can rapidly germinate in response to specific germinants

that signal conditions conducive for vegetative growth (Setlow, 2014; Sinai *et al.*, 2015; Sturm and Dworkin, 2015).

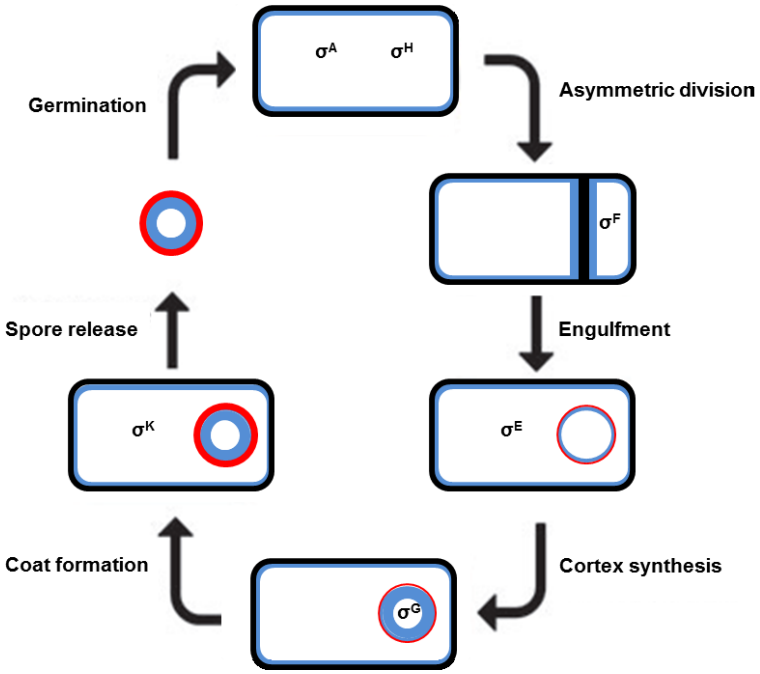


Fig.2. Steps in the sporulation pathway of *Bacillus subtilis*. Upon starvation, *B. subtilis* cells slow down in growth and initiate the sporulation process. Asymmetric division results in formation of a smaller forespore and a larger mother cell. Both cells follow distinct developmental programs of gene expression driven by stage- and compartment-specific transcription factors linked to each other through cell-cell signaling pathways. This ensures that gene expression in one cell is kept in register with gene expression in the other cell. Upon maturation of the developing forespore, lysis of the mother cell releases a highly stress-resistant endospore into the environment As soon as the dormant spore encounters appropriate environmental stimuli; it initiates germination resulting in reentry to vegetative growth. Modified from Higgins and Dworkin (Higgins and Dworkin, 2012).

Entry into the sporulation pathway is governed by two transcription factors: the master response regulator Spo0A and the alternative sigma factor σ^H. These sporulation-inducing transcription factors are under complex and interconnected regulatory control including several positive and negative feedback-loops and guide expression of hundreds of genes at the onset of differentiation (Chastanet and Losick, 2011; Grossman, 1995; Hoch, 1993; Phillips and Strauch, 2002; Sonenshein, 2000). As soon as a specific threshold level of activated Spo0A~P necessary to induce sporulation is achieved the cell enters the developmental program reaching a point of no return (Molle *et al.*, 2003). Hence, *B. subtilis* cells need to assure that the dwindling nutritional and energetic resources available to it, and the environmental circumstances, will allow completion of the sporulation process since otherwise, both, the mother cell will lyse but no spore has been formed. Indeed, previous work demonstrated that starving *B.*

subtilis cells simultaneously facing osmotic stress are impaired in spore formation (Kunst and Rapoport, 1995; Ruzal *et al*., 1998; Ruzal and Sanchez-Rivas, 1998). Based on these observations, we found in a side project of this dissertation that starving *B. subtilis* cells prevent spore formation when simultaneously coping with osmotic stress thereby avoiding death during cellular differentiation (Publication 7, Appendix II) (Widderich *et al*., 2016b). This is an interesting example how cells coping with multiple stresses, a situation they also frequently encounter in their natural habitats, ensure elevated likelihood of their survival and that of the community by integrating signals obtained by monitoring different cellular and environmental parameters.

III. General stress response

The general stress response of microorganisms can be triggered by a variety of different stress conditions. It is characterized by plenty modifications in cellular physiology and even morphology and should guarantee that cellular damage is being prevented rather than repaired. As a consequence, the general stress resistance renders microorganisms broadly stress resistant, even against stresses they have not experienced elsewhere before, and enhances survival of the community (Storz and Hengge-Aronis, 2000). Newly arising stresses usually result in accumulation of an alternative sigma factor (e.g. RpoS in *Escherichia coli*, SigB in *Bacillus subtilis*) that acts as the master regulator of the general stress response and coordinates the expression of a multitude of genes (Hecker *et al*., 2007; Hengge-Aronis, 2000; Hengge-Aronis, 2002; Price, 2000; Storz and Hengge-Aronis, 2000; Francez-Charlot *et al*., 2015). Their corresponding gene products not only confer increased stress tolerance but may also modify cellular metabolism and mediate structural and morphological rearrangements. Loss of the alternative sigma factor underlines its important role in the general stress response since its deletion results in strong sensitivity to multiple stresses and concomitantly deficiency in survival (Hengge-Aronis, 2000; Price, 2000).
In *B. subtilis*, the general stress response is controlled by σ^B, whose transcription is activated by several growth-limiting environmental and metabolic stresses (Hecker *et al*., 2007). The SigB-controlled regulon comprises about 252 potential genes and their expression can amount about approximately 30 % of the newly synthesized proteins when facing severe stress (Nannapaneni *et al*., 2012; Price, 2000). A good studied example of the general stress response in *B. subtilis* is the contribution of σ^B to osmotic stress resistance. In particular, it has been demonstrated that the expression of transport systems involved in osmo-adaptation is up-regulated in dependence of σ^B in

response to a salt-shock (Nicolas *et al*., 2012; Steil *et al*., 2005). Moreover, mutants defective in members of the σ^B-controlled general stress regulon display an increased salt-sensitive phenotype underlining its important role in osmostress adaptation (Hoper *et al*., 2005), however, the specific role of most member genes of the σ^B-controlled regulon is still not known. Although salt-shock is one of the strongest inducers of the general stress response in *B. subtilis* (Nannapaneni *et al*., 2012), induction of σ^B-controlled genes by this environmental cue is only transiently and short-lived (Young *et al*., 2013). In addition, it has been demonstrated that the general stress regulon is not induced in cells that experience chronic exposure to high salinity (Spiegelhalter and Bremer, 1998). Consequently, *B. subtilis* and other microorganisms have developed further strategies, the specific stress responses, in order to cope with long-lasting stressful circumstances (Storz and Hengge-Aronis, 2000).

IV. Osmotic stress response

In contrast to the general stress response and to cellular differentiation processes, many microorganisms have also evolved a considerable number of stress responses that are induced under specific stressful circumstances. These responses actually repair occurring damage and/or eliminate the detrimental agent (Storz and Hengge-Aronis, 2000). The osmotic stress response is one of the best studied single-induced stress responses in microorganisms. Changing osmolarity is, due to its impact on almost all microorganisms, one of the most important parameters of an ecosystem (Csonka and Epstein, 1996; Kempf and Bremer, 1998; Lucht and Bremer, 1994). Many microorganisms, such as the ubiquitously distributed soil-dwelling bacterium *B. subtilis* (Mandic-Mulec *et al*., 2015), frequently face changing osmolarity/salinity due to fluctuations in water availability caused by either flood or drought of a habitat (Bremer and Krämer, 2000; Pittelkow and Bremer, 2011; Booth and Blount, 2012).

The dilution of osmotically active substances by e.g. heavy rains results in the establishment of a hypoosmotic milieu. Hence, the resulting intracellular ionic strength is much higher than its surrounding and triggers the influx of water into the microbial cell alongside the osmotic gradient (Csonka and Epstein, 1996; Lucht and Bremer, 1994). This inevitably causes an increase in the cellular internal pressure that finally leads to the burst of the cell (Booth *et al*., 2007). In contrast, long-lasting droughts cause the opposite effect. The desiccation of a natural habitat by different abiotic environmental factors, such as solar radiation and heat, leads to an increase in external salinity and osmolarity. Establishment of a hyperosmotic milieu by the enrichment of osmotically active compounds triggers the efflux of water across the

cytoplasmic membrane out of a microbial cell (Bremer and Krämer, 2000; Pittelkow and Bremer, 2011). This water efflux impairs the maintenance of the cellular vital functions after the onset of plasmolysis that results from dehydration of the cytoplasm (Korber *et al*., 1996). These specified variations negatively affect the cellular internal pressure, the turgor. The permeability of the cytoplasmic membrane for water, as well as a considerable osmotic potential of the cellular cytoplasm, establish the turgor, an intracellular hydrostatic pressure whose maintenance within physiologically appropriate boundaries is considered to be critical for cell integrity and division (Koch, 1983; Booth and Blount, 2012; Booth *et al*., 2007; Bremer and Krämer, 2000; Wood, 2011). According to the 'surface stress theory', a positive turgor is essential for the synthesis of cell wall material and consequently ensures growth and proliferation of a cell (Koch, 1983). Therefore, microbial cells need to actively adapt to changes in their environment in order to avoid a collapse of turgor and to guarantee survival (Miller and Wood, 1996; Pittelkow and Bremer, 2011). Since microorganisms are not able to actively pump water across their membrane, they have developed other strategies to indirectly regulate the water balance and to adapt to changing osmolarity (Galinski, 1995; Pittelkow and Bremer, 2011)

To counteract the efflux of water under hyperosmotic conditions and to maintain cell integrity, basically two different strategies have been evolved by microorganisms (Pittelkow and Bremer, 2011). The first is called the 'salt in' -strategy and is based on the intracellular accumulation of large quantities of inorganic ions, mainly potassium (K^+) and chloride (Cl^-), and the concomitant extrusion of cytotoxic sodium (Na^+) ions (Galinski and Trüper, 1994; Ventosa *et al*., 1998). However, occurrence of these microbes is mostly restricted to high saline habitats, since their lifestyle is fully adapted to these extreme environments and their protein biochemistry and molecular biology is dependent on a high intracellular ionic strength. For this reason these microorganisms are usually entitled as 'obligate halophiles' (Coquelle and Glover, 2010; Grant, 2004; Rhodes *et al*., 2010). The second strategy – the 'salt out'-strategy – is more widely distributed. Here, low-molecular organic compounds, so-called compatible solutes, are enriched in the cytoplasm instead of inorganic ions. These osmolytes are accumulated up to molar concentrations by the uptake via specialized transport systems or by *de novo* -synthesis preventing the diffusion of water under hyperosmotic conditions and thereby avert impending plasmolysis (Bremer and Krämer, 2000; Csonka and Epstein, 1996; Kempf and Bremer, 1998). Microorganisms using this strategy are able to cope with increased salinities in their environment but do not rely on it. Hence, these microbes are entitled as 'halotolerants'.

When microorganisms face a hypoosmotic shock as for example caused by heavy rainfall, they also need to be able to quickly reduce the concentration of osmotically active substances. Since degradation of these compounds normally takes up a considerable amount of time, many microorganisms have evolved systems that mediate extrusion of osmotically active substances in a timely manner. Most microorganisms possess pore-forming protein-complexes, so called mechanosensitive channels. These channel-forming proteins localize in the cytoplasmic membrane and are transiently opened by increasing membrane tension as caused by elevating turgor pressure (Haswell *et al*., 2011). This allows extrusion of large amounts of solutes in a short manner, but also non-selectively. Therefore, these channels function as safety-release valves before the cell bursts (Booth *et al*., 2007; Hoffmann *et al*., 2008; Kung *et al*., 2010). Many microorganisms possess different types of mechanosensitive channels (MscS, MscM, MscL) with different pore diameters that allow a graded response under these challenging conditions.

V. Compatible solutes

Compatible solutes are low-molecular, highly soluble and osmotically active compounds that can be accumulated to very high intracellular concentrations without interfering with the cellular metabolism (Kempf and Bremer, 1998; Wood *et al*., 2001; Kurz, 2008). Moreover, these osmolytes are highly water-soluble and possess the ability to stabilize proteins and other macromolecules *in vitro* as well as *in vivo* (Arora *et al*., 2004; Canovas *et al*., 1999; Lippert and Galinski, 1992). Due to these properties many potent compatible solutes are also designated as *chemical chaperones* (Tatzelt *et al*., 1996). The preferential exclusion model proposed by Arakawa and Timasheff is usually used to explain this property of compatible solutes (Arakawa and Timasheff, 1985). In general, this term is based on the osmophobic interplay that occurs between the osmolytes and the backbone of proteins. Due to these interactions, the organic solutes are excluded from the hydration shell of macromolecules, resulting in the concomitant reduction of the captured volume by the macromolecules to a minimum and thereby stabilizing their native conformation (Arakawa and Timasheff, 1985; Bolen and Baskakov, 2001).

Chemically, compatible solutes are divided into four major groups: i) sugars and polyoles (e.g. trehalose and glycerol), ii) trimethylammonia-compounds (e.g. glycine betaine and carnitine), iii) methylsulfonio-compounds and sulfate esters (e.g. dimethylsulfoniopropionate and choline-O-sulfate), as well as iv) amino acids (e.g. proline and glutamate) and their derivatives (e.g. ectoine and hydroxyectoine) (Bursy *et*

al., 2007; Inbar and Lapidot, 1988; Kempf and Bremer, 1998; Miller and Wood, 1996; Wood *et al*., 2001). Some of the most important and abundant representatives of compatible solutes are shown in Figure 3.

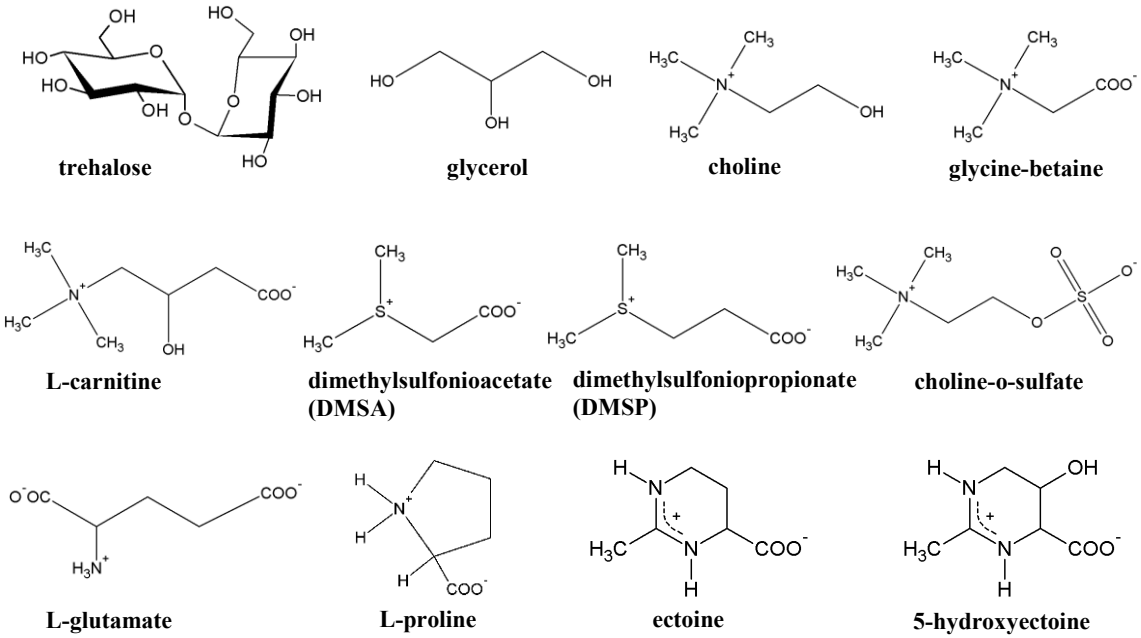


Fig. 3. Selection of important representatives of compatible solutes. Chemical structures of selected compatible solutes are depicted. Ectoine and its derivative 5-hydroxyectoine are in spotlight of the present dissertation. Image is adapted and modified from Widderich (Widderich, 2012).

As already mentioned above, compatible solutes can be either synthesized *de novo* or can be taken up from the environment (Kempf and Bremer, 1998). The latter is generally preferred by microorganisms since the uptake of these osmolytes is in principal energetically more favorable (Oren, 1999). Beside their osmoprotectant properties, some compatible solutes also protect microorganisms against temperature stress (Bursy *et al*., 2008; Garcia-Estepa *et al*., 2006; Holtmann and Bremer, 2004; Kempf and Bremer, 1998; Kuhlmann *et al*., 2011; Bashir *et al*., 2014a; Bashir *et al*., 2014b). For example, the potent osmolyte glycine betaine allows *B. subtilis* cells to grow at 13° C, a temperature that is usually too cold for this organism to proliferate in a chemically defined medium (Hoffmann and Bremer, 2011). Moreover, compatible solutes have also been shown to increase enzyme activities and to protect cells against oxidative stress (Cyplik *et al*., 2012; Andersson *et al*., 2000; Smirnoff and Cumbes, 1989). Simply put, this considerable list of protective properties also explains the fact that distribution of compatible solutes is not only restricted to prokaryotes, such as *Bacteria* and *Archaea* (da Costa *et al*., 1998; Empadinhas and da Costa, 2008; Kempf

and Bremer, 1998). In fact, a great deal of the biological input of compatible solutes into the environment also occurs by plants, fungi and animals (Bremer, 2002; Csonka, 1989).

The occurrence and distribution of compatible solutes is quite well understood in various halophilic and halo-tolerant microorganisms. The utilization of compatible solutes, at least in many *Bacteria*, seems thereby to occasionally correlate with their position within the phylogenetic *tree of life* (Oren, 2008). For instance, the ability to synthesize ectoines seems to be mainly restricted to halophilic and halo-tolerant *Proteobacteria*, *Actinobacteria* and *Firmicutes* (Pastor *et al*., 2010; Pittelkow, 2011), whereas the synthesis of trehalose and glutamate in response to osmotic stress, as well as the uptake and accumulation of glycine betaine via specific transport systems, appears to be more common and even wider distributed among microorganisms (da Costa *et al*., 1998; Pittelkow and Bremer, 2011).

VI. The compatible solute ectoine and its derivative hydroxyectoine

One of the most abundant compatible solutes in nature is the tetrahydropyrimidine ectoine [(*S*)-2-methyl-1,4,5,6-tetra-hydropyrimidine-4-carboxylate]. In the first instance, ectoine has been biochemically assigned to various groups of chemical compounds. That way, ectoine has been described for example as a heterocyclic amino acid, a *N*-acetylic di-amino acid and a hydrogenated pyrimidine derivative (Galinski, 1993; Peters *et al*., 1990). It is synthesized by a large number of microorganisms in response to osmotic and temperature stress (Bursy *et al*., 2008; Bursy *et al*., 2007; Grant, 2004; Kuhlmann *et al*., 2008a; Kuhlmann *et al*., 2011; Ono *et al*., 1999; Ventosa *et al*., 1998) and has been initially identified and described as a compatible solute in the gram-negative, halophilic and phototrophic sulfurbacterium *Ectothiorhodospira halochloris* (Galinski *et al*., 1985). Shortly thereafter, 5-hydroxyectoine [(*S,S*)-*β*-2-methyl-5-hydroxy-1,4,5,6-tetrahydropyrimidine-4-carboxylate] has been discovered in the soil-living bacterium *Streptomyces parvulus* and has been identified as a compatible solute as well (Inbar and Lapidot, 1988).

Hydroxyectoine is a hydroxylated derivative of ectoine (Bursy *et al*., 2007; Inbar *et al*., 1993; Inbar and Lapidot, 1988). Hence, both compounds are closely related and only differ in a hydroxyl-group positioned at the C5-atom of the pyrimidine ring structure of hydroxyectoine (Fig. 4). Both, ectoine and hydroxyectoine serve as excellent osmoprotectants (Borges *et al*., 2002; Schiraldi *et al*., 2006). Consequently, disruptions of their corresponding biosynthetic genes result in strongly salt-sensitive phenotypes (Goller *et al*., 1998). Due to their benign nature, they also possess the ability to

stabilize the cytoplasmic membrane, protect macromolecules, such as the DNA, from the detrimental effects of UV-radiation and prevent denaturation of proteins under temperature stress (Bursy *et al*., 2008; Fletcher and Csonka, 1998; Jebbar *et al*., 1992; Kolp *et al*., 2006; Kuhlmann *et al*., 2008a; Malin and Lapidot, 1996; Roychoudhury *et al*., 2012; Vargas *et al*., 2008).

Despite their very close relatedness in chemical structure, both compounds vary in some aspects with respect to their impact and stabilizing properties on macromolecules. For instance, ectoine decreases the melting temperature of double-stranded DNA, whereas the presence of hydroxyectoine results in an increase of the melting temperature (Kurz, 2008; Schnoor *et al*., 2004). Moreover, differences in their stabilizing properties on macromolecules have been reported in *in vitro* assays, being hydroxyectoine quite often the more potent stress protectant and the more effective stabilizing agent (Borges *et al*., 2002; Harishchandra *et al*., 2010; Lippert and Galinski, 1992; Manzanera *et al*., 2004), in particular with respect to temperature and desiccation stress (Bursy *et al*., 2008; Garcia-Estepa *et al*., 2006; Manzanera *et al*., 2002; Tanne *et al*., 2014). These superior properties are attributed to the additional hydroxyl-group at the C-atom of the pyrimidine ring. Accordingly, the presence of this additional hydroxyl-group within the hydroxyectoine molecule allows stronger interaction with the polar groups of proteins and other macromolecules and might thus be capable to replace missing water molecules (Bursy *et al*., 2007; da Costa *et al*., 1998; Galinski, 1993). In this context, it has also been demonstrated that these stronger intermolecular H-bonds with the OH-group make hydroxyectoine, in contrast to ectoine, a good glass-forming compound (Tanne *et al*., 2014).

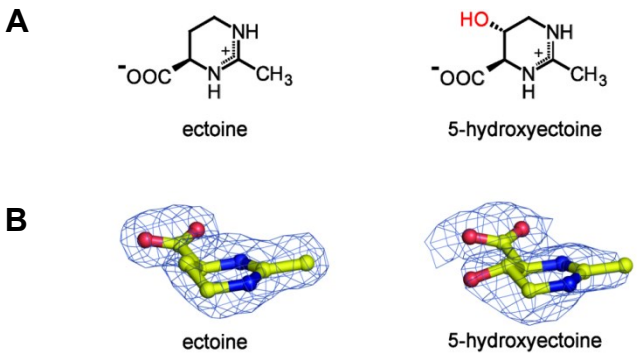


Fig. 4. Chemical structures of ectoine and 5-hydroxyectoine. (A) Projection structure of ectoine and its hydroxylated derivative 5-hydroxyectoine. (B) Ball-and-stick representation of ectoine and 5-hydroxyectoine along with the electron density around these compounds in the crystal structure of the EhuB ectoine/5-hydroxyectoine binding protein of *Sinorhizobium meliloti*; these data were extracted from PDB files 2Q88 and 2Q89 for the EhuB:ectoine and EhuB:5-hydroxyectoine complexes (Hanekop *et al*., 2007).

VII. Genetic organization of ectoine and hydroxyectoine biosynthetic genes

Genetic analysis of ectoine- and hydroxyectoine-producing microorganisms revealed that the corresponding ectoine biosynthetic genes are usually encoded in a gene cluster (*ectABC*) and are typically transcribed as an operon from a single, osmotically inducible promoter (Bursy, 2005; Pittelkow, 2011) (Fig. 5). In this context, it should be mentioned that the organization of the *ect*-cluster (*ectABC*) does not reflect the order of enzymatic steps (EctB, EctA, EctC) in the ectoine biosynthesis pathway. The previous cited studies also showed a high degree of relatedness for both the ectoine biosynthetic enzymes as well as their corresponding genes suggesting that the ectoine biosynthesis pathway is evolutionary highly conserved (Goller *et al*., 1998; Kuhlmann and Bremer, 2002; Louis and Galinski, 1997; Ono *et al*., 1999). Disruption of the ectoine biosynthetic genes results in a strong salt- and osmo-sensitive phenotype demonstrating the important role of ectoines in bacterial osmo-adaptation (Goller *et al*., 1998).

A sub-group of the ectoine-producing microorganisms is able to convert ectoine into its hydroxylated derivative 5-hydroxyectoine. The genetic information responsible for this reaction is encoded by the *ectD* gene (Bursy *et al*., 2008; Bursy *et al*., 2007; Galinski and Trüper, 1994; Kuhlmann and Bremer, 2002; Prabhu *et al*., 2004). This gene can be found in either of two genetic contexts: i) it may be co-transcribed with the *ectABC* operon and is usually located downstream of the *ectC* gene (Bursy *et al*., 2008; Mustakhimov *et al*., 2010), or ii) it may be located at another position in the genome (Bursy *et al*., 2007; Pittelkow, 2011).

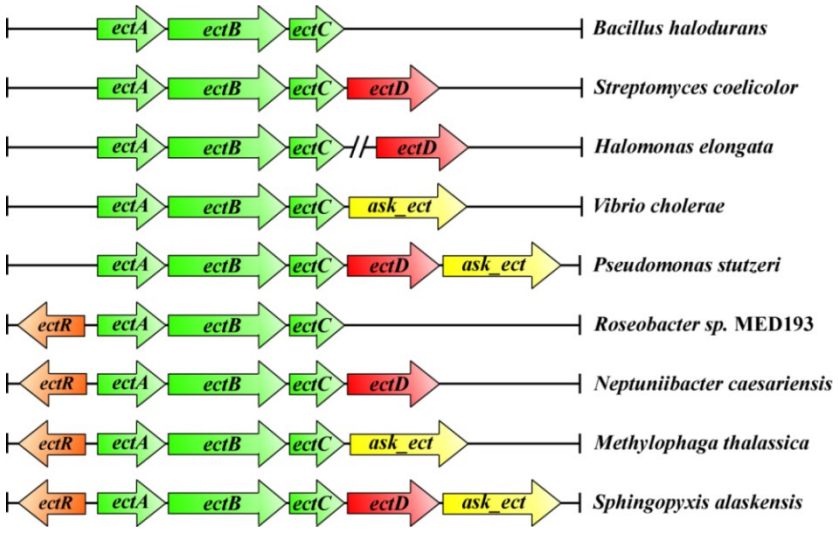


Fig. 5. Genetic organization of the ectoine biosynthetic genes. Shown are the most common organizational types of gene clusters for ectoine/5-hydroxyectoine biosynthesis present in ectoine/5-hydroxyectoine producers (including examples). Image is adapted from Widderich (Widderich, 2012).

VIII. Peculiarities of the ectoine biosynthesis gene cluster

As detailed above, the ectoine biosynthetic genes are typically encoded in gene cluster. However, in a subset of microorganisms the *ect*-cluster exhibits different special features and/or peculiarities. For instance, variants of the *ect*-gene cluster have been described comprising additional genes (Pittelkow, 2011; Widderich, 2012). Some clusters contain potential *mfs*-genes that presumably encode transport systems of the major facilitator superfamily that might be responsible for the import or export of ectoines. In the gene neighborhood of other *ect*-clusters genes have been identified that encode potential ABC transport systems. Bioinformatic analysis of these genes revealed high similarity to already characterized ectoine uptake systems, such as EhuB from *Sinorhizobium meliloti* (Pittelkow, 2011; Widderich, 2012).

In a substantial number of ectoine producers an additional *ask_ect* gene can be found as part of the ectoine biosynthesis gene cluster encoding an aspartokinase (Stöveken *et al*., 2011). This enzyme catalyzes the reaction from L-aspartate, a precursor substrate of the ectoine biosynthesis pathway, to aspartate-β-phosphate, which is, in a further enzymatic step, converted into aspartate-β-semialdehyde by the aspartate-β-semialdehyde-dehydrogenase (Asd) (Fig. 6). In many microorganisms, aspartate-β-semialdehyde serves as a central hub in a diversity of different metabolic pathways (Lo *et al*., 2009). For instance, it is the starting substance for the biosynthesis of the amino acids methionine, threonine and lysine (Cohen and Saint-Girons, 1987). Moreover, it is the precursor molecule for the production of dipicolinic acid, which plays a central role in spores of gram-positive bacteria, and of *meso*-diaminopimelic acid, which is used by most microorganisms to crosslink peptidoglycan during cell wall synthesis (Cohen and Saint-Girons, 1987). But most notably in the context of this dissertation is that aspartate-β-semialdehyde also serves as the substrate for the ectoine biosynthesis pathway (Fig. 6 and 7).

Phosphorylation of L-aspartate by the aspartokinase is driven through consumption of ATP and is therefore energy-demanding. Hence, to avoid loss of energy-equivalents, activity of aspartokinases and transcription of their corresponding genes is, for the purpose of metabolic pathways, in general strictly feedback-regulated and under stringent transcriptional control (Lo *et al*., 2009). However, the Ask_ect protein encoded in the ectoine gene cluster of *Pseudomonas stutzeri* shows no such strict feedback-inhibition (Stöveken *et al*., 2011). It is predicted that the development of such an enzyme, whose genetic information is associated to the ectoine biosynthetic gene cluster, might ensure sufficient availability of precursor molecules for ectoine biosynthesis under osmotic stress conditions and thereby provides an improved

adaptation process to these challenging conditions (Bestvater *et al*., 2008; Stöveken *et al*., 2011).

In another group of ectoine producing microorganisms a regulatory gene (*ectR*) has been described, whose gene product EctR is a member of the widely distributed group of MarR transcriptional regulators, but forms a distinct sub-group within this superfamily. It has been shown that EctR serves as a repressor for the *ectABC-ask_ect* operon (Mustakhimov *et al*., 2012; Mustakhimov *et al*., 2010), however, elevated transcription of the ectoine biosynthetic genes in an *ectR* mutant remained osmotically inducible (Mustakhimov *et al*., 2010). In *Methylomicrobium alcaliphilum*, the EctR operator overlaps the -10 sequence of the *ect* promoter and EctR might also regulate the expression of its own structural gene; but the latter regulatory feature does not always seem to exist (Mustakhimov *et al*., 2012). Moreover, it is currently still unknown which environmental or cellular cues dictate the binding to or the release of EctR from its operator sequence.

In those ectoine/5-hydroxyectoine producers that possess an *ectR* gene, it is mostly transcribed in a divergent orientation from the ectoine/5-hydroxyectoine biosynthetic gene cluster (Fig. 5).

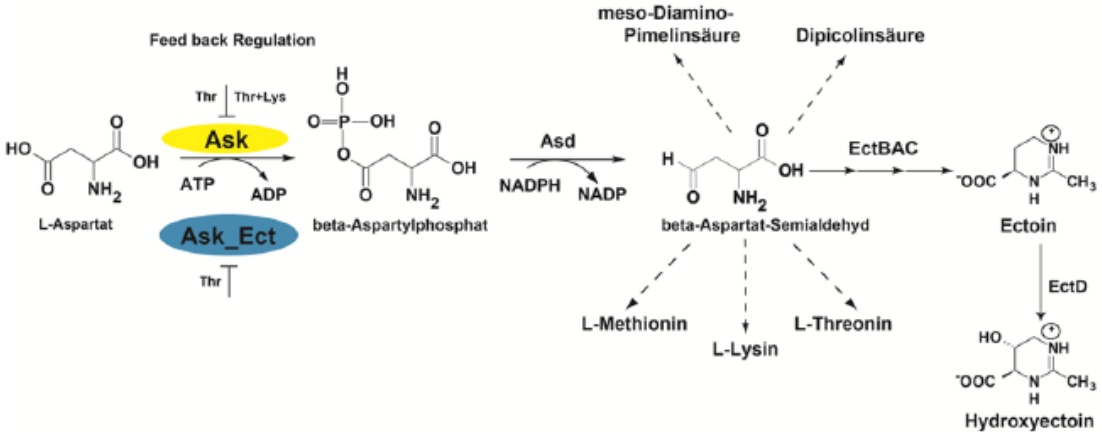


Fig. 6. Scheme of the role of aspartokinases in cellular metabolism. Asprtokinases catalyze the phosphorylation of L-aspartate to yield aspartyl-β-phosphate which is subsequently converted to aspartate-β-semialdehyde by the Asd enzyme. Aspartate-β-semialdehyde is the precursor molecule for many microbial metabolic pathways including ectoine biosynthesis. Image is adapted from Pittelkow (Pittelkow, 2011).

IX. Ectoine and hydroxyectoine uptake systems

In order to accumulate compatible solutes, such as the chemical chaperones ectoine and its derivative hydroxyectoine, these compounds need to be either synthesized *de novo* as described in detail above or to be taken up by specialized transport systems.

For instance, different types of uptake systems are known in the gram-positive bacterium *B. subtilis* that channel compatible solutes into the cytoplasm under osmotic stress conditions (Kappes *et al*., 1996; Kappes *et al*., 1999; Kempf and Bremer, 1995). This is accomplished by secondary active transporters, such as OpuD and OpuE, and by primary active ABC (ATP binding cassette)-transporters, such as OpuA, OpuB and OpuC, which are driven by the consumption of ATP. In case of OpuC, it has already been shown that it is able to channel ectoine into the cytoplasm. However, the binding affinity to ectoine by this transporter is quite low, when compared to the other solutes that can be taken up by OpuC (Jebbar *et al*., 1997; Du *et al*., 2011).

Numerous microorganisms thus evolved high affinity uptake systems for ectoine and hydroxyectoine. These belong, quite often, also to the already mentioned class of ABC transporters, but other types, in example TRAP (tripartite ATP-independent periplasmic)-transporters have been identified as well (Hanekop *et al*., 2007; Kuhlmann *et al*., 2008b; Lecher *et al*., 2009). In both cases, specific substrate binding proteins mediate the high affinity binding of ectoines. The ABC transport system Ehu (ectoine and hydroxyectoine uptake) from *Sinorhizobium meliloti* is one example for such a high affinity ectoine transporter and a crystal structure of its binding protein EhuB has already been solved (Hanekop *et al*., 2007). This crystal structure unambiguously revealed the amino acids essential for the binding of ectoines. A similar binding motif has also been observed in the crystal structures of the substrate binding protein UehA from *Silicibacter pomeroyii* and TeaA from *Halomonas elongata* (Kuhlmann *et al*., 2008b; Lecher *et al*., 2009). In the meantime, it is known that a multiplicity of microorganisms possesses the capability to accumulate ectoines via specialized uptake systems in order to protect themselves against the detrimental effects caused by high environmental osmolarity (Galinski *et al*., 1985; Galinski and Trüper, 1994; Inbar and Lapidot, 1988; Kempf and Bremer, 1995). A well-studied example is EctT, a member of the betaine-choline-carnitine transporter (BCCT)-type carrier family, which mediates the import of ectoine and hydroxyectoine as protectants against osmotic and cold stress in *Virgibacillus pantothenticus* (Kuhlmann *et al*., 2011).

X. Biosynthesis of ectoines

The reaction steps for the synthesis of ectoine have been initially deciphered in the microorganisms *Streptomyces parvulus* und *Halomonas elongata* (Inbar and Lapidot, 1988; Ono *et al*., 1999; Peters *et al*., 1990). Ectoine is synthesized from L-aspartate-β-semialdehyde, a central intermediate of the amino acid-metabolism (Louis and Galinski, 1997), via a three-stepped enzymatic reaction (Inbar and Lapidot, 1988; Ono

et al., 1999; Peters *et al*., 1990) (Fig. 7). Synthesis of ectoine occurs in the first step from an amino transfer from the co-substrate L-glutamate to L-aspartate-β-semialdehyde to form L-2,4-diaminobutyrate and the by-product 2-oxoglutarate. This reaction is catalyzed by the L-2,4-diaminobutyrate transaminase EctB. In the next step, L-2,4-diaminobutyrate is acetylated at the γ-amino group to yield *N*-γ-acetyl-2,4-diaminobutyrate via a transfer of the acetyl-group from the co-substrate acetyl-CoA by the 2,4-diaminobutyrate acetyltransferase EctA. This intermediate is in the last step of ectoine biosynthesis subsequently dehydrated by the key enzyme of ectoine biosynthesis, the ectoine synthase EctC. The elimination of a water molecule from *N*-γ-acetyl-2,4-diaminobutyrate results in ring closure and formation of the cyclic ectoine molecule (Kuhlmann and Bremer, 2002; Louis and Galinski, 1997). A considerable number of ectoine producers is also able to convert ectoine into its hydroxylated derivative 5-hydroxyectoine. The enzyme catalyzing this reaction is the ectoine hydroxylase EctD (Bursy, 2005; Bursy *et al*., 2007). In this reaction step, one molecular oxygen atom of the co-substrate O_2 is transferred to the C5-atom of the pyrimidine ring generating the hydroxyl-group. The other molecular oxygen atom is transferred in a side-reaction to the second co-substrate 2-oxoglutarate resulting in the elimination of the carboxyl-group, which is released as CO_2, and the formation of succinate. In summary, the ectoine hydroxylase catalyzes the reaction of ectoine, O_2 and 2-oxoglutarate to yield 5-hydroxyectoine, carbon dioxide and succinate. A detailed schematic view of the ectoine and hydroxyectoine biosynthesis pathway is shown in Fig. 7.

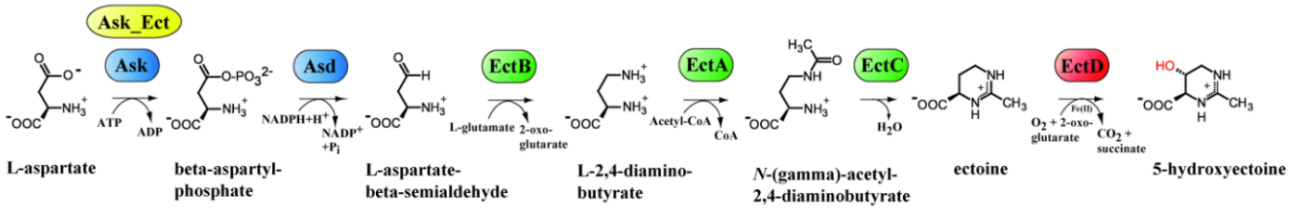


Fig. 7. Pathway for the synthesis of ectoines. The shown pathway for the synthesis of ectoines is based on data published by Peters *et al*., Ono *et al*. and Bursy *et al*. (Bursy *et al*., 2007; Ono *et al*., 1999; Peters *et al*., 1990). Ectoine-biosynthesis is catalyzed by a three-stepped enzymatic reaction. In the first step, the L-2,4-diaminobutyrate-transaminase (EctB) catalyzes the reaction from L-aspartate-β-semialdehyde to L-2,4-diaminobutyrate, which afterwards is directly converted to Nγ-acetyl-L-2,4-diaminobutyrate by the L-2,4-diaminobutyrate-Nγ-acetyltransferase (EctA). In the last reaction step, the condensation of Nγ-acetyl-L-2,4-diaminobutyrate by the ectoine synthase (EctC) results in ring closure and ectoine formation. In a substantial number of ectoine producers, this cyclic compound can be additionally be hydroxylated at the C5-position of the pyrimidine ring to form its derivative 5-hydroxyectoine. The enzyme responsible for this reaction is the ectoine-hydroxylase (EctD).

XI. EctC and EctD, the key enzymes in ectoine and hydroxyectoine biosynthesis

The ectoine synthase EctC catalyzes the cyclocondensation reaction of *N*-γ-acetyl-L-2,4-diaminobutyric acid to yield ectoine (Kuhlmann and Bremer, 2002; Louis and Galinski, 1997; Ono *et al*., 1999), which serves as the substrate for the ectoine hydroxylase EctD to produce its derivative 5-hydroxyectoine (Bursy *et al*., 2007). While the EctA and EctB enzymes have counterparts in various microbial metabolic pathways that are highly related in their amino acid sequence to EctA and EctB, this is not true for the EctC and EctD proteins. Hence, these enzymes can be considered as the key enzymes in ectoine and hydroxyectoine biosynthesis, respectively.

Based on bioinformatic approaches, the ectoine synthase belongs to the cupin superfamily, a large group of proteins having a conserved ß-barrel structure scaffold (Dunwell *et al*., 2001). This protein family is known to perform a variety of enzymatic and non-enzymatic functions (Dunwell *et al*., 2001; Agarwal *et al*., 2009; Uberto and Moomaw, 2013). The biochemical properties of a few ectoine synthases deriving from *H. elongata*, *M. alcaliphilum* and *Acidiphilium cryptum* have already been reported in the literature (Moritz *et al*., 2015; Ono *et al*., 1999; Reshetnikov *et al*., 2011; Reshetnikov *et al*., 2006). Each of the studies revealed that EctC catalyzes the cyclization of *N*-γ-acetyl-L-2,4-diaminobutyric acid (*N*-γ-ADABA), the reaction product of the 2,4-diaminobutyrate acetyltransferase (EctA), to yield ectoine by concomitantly releasing a water molecule. Furthermore, it has been reported that the ectoine synthase is also able to catalyze, in a minor side reaction, the hydrolysis of ectoine. It can thereby also use ectoine derivatives with reduced or expanded ring structures (Moritz *et al*., 2015; Witt *et al*., 2011). Moreover, some data suggest that EctC can also catalyze the formation of the synthetic compatible solute 5-amino-3,4-dihydro-2H-pyrrole-2-carboxylate (ADPC) through the cyclic condensation of two glutamine molecules (Witt *et al*., 2011). Although some progress has been made with respect to the biochemical characterization of the ectoine synthase, our understanding of the structure and architecture of the active site of this enzyme is still incomplete, since its biochemical properties have only been poorly studied and a crystal structure of this enzyme has still not been solved.

The ectoine hydroxylase belongs to the superfamily of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases (Bursy *et al*., 2007; Reuter *et al*., 2010). This family of proteins comprises versatile biocatalysts not only carrying out hydroxylation reactions but also demethylations, desaturations, cyclizations, ring expansions, epimerizations, and halogenations (Aik *et al*., 2012; Hangasky *et al*., 2013;

Hausinger, 2004; Kundu, 2012; Wong *et al*., 2013). Although the type of substrates used by these dioxygenases varies considerably, common enzyme reaction mechanisms are observed. The majority of these enzymes couples a two-electron oxidation of their substrates with the reaction of oxygen and 2-oxoglutarate (Aik *et al*., 2012; Grzyska *et al*., 2010; Hangasky *et al*., 2013), as it is also the case for the ectoine hydroxylase (Bursy *et al*., 2007; Reuter *et al*., 2010). A crystal structure of the ectoine hydroxylase from the moderate halophile *Virgibacillus salexigens* has already been solved and deposited in the RCSB Protein Data Bank under accession number 3EMR (Reuter *et al*., 2010). This structure corresponded to a monomer and its core consisted of a double-stranded β-helix (DSBH), flanked and stabilized by several α-helices (Fig. 8). The DSHB is also often referred to as jelly roll fold and is formed by two four-stranded anti-parallel β-sheets arranged in form of a β-sandwich (Aik *et al*., 2012; Hangasky *et al*., 2013; Hausinger, 2004; Kundu, 2012).

The solved EctD structure from the moderate halophile *Virgibacillus salexigens* also unambiguously revealed the positioning of a mononuclear iron center within the active site of the enzyme. However, it contained neither the co-substrate 2-oxoglutarate nor the substrate ectoine (Reuter *et al*., 2010). In the active site of ectoine hydroxylase, the iron co-factor is coordinated by the side chains of His-146, Asp-148, and His-248 that protrude into the ligand-binding cavity of the *Vs*EctD protein and form a HxD/E…H motif, the so called 2-His-1-carboxylate facial triad that is conserved among the superfamily of non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases (Aik *et al*., 2012; Hangasky *et al*., 2013; Hausinger, 2004; Knauer *et al*., 2012; Prescott and Lloyd, 2000; Straganz and Nidetzky, 2006). Moreover, the iron-coordinating residues His-146 and Asp-148 are part of a strictly conserved signature sequence motif for EctD-type hydroxylases (extending from Phe143 to Pro159) (Bursy *et al*., 2007; Reuter *et al*., 2010). This 17 amino acids comprising consensus sequence spans an extended α-helix and a linked short β-strand within the EctD crystal structure and form one site of the presumed active site of the ectoine hydroxylase (Reuter *et al*., 2010). Initial biochemical characterization of EctD-type proteins has also been carried out in previous studies in combination with site-directed mutagenesis experiments (Bursy *et al*., 2007; Pittelkow, 2011), but trials to solve the crystal structure with all bound ligands failed (Reuter *et al*., 2010). Hence, in-depth knowledge of the architecture and biochemistry of the ectoine hydroxylase was still missing at the onset of my PhD studies.

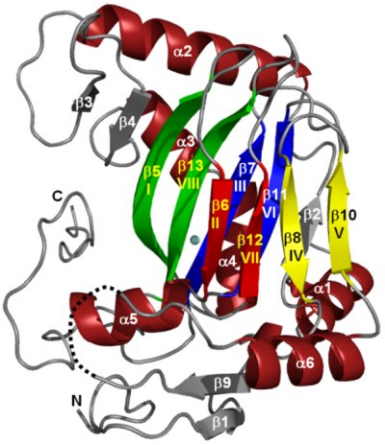


Fig. 8. Crystal structure of the ectoine hydroxylase EctD from *Virgibacillus salexigens*. The crystal structure of the ectoine hydroxylase EctD from *V. salexigens* has been solved by X-ray structure analysis with a resolution of 1.8 Å. This figure shows the solved crystal structure of the EctD enzyme. The enzymes α-helices and β-sheets are depicted in a numerical order and the iron(II)-ligand is highlighted as a blue ball between the beta-sheets β-II and β-VII. Image adapted from Reuter *et al.* (Reuter *et al*., 2010).

XII. Ectoines in industry

Due to their highly stabilizing effects on biological macromolecules and entire cells, both ectoine and hydroxyectoine have spurred considerable biotechnological interest and are used in a variety of different products (Graf *et al*., 2008; Kunte *et al*., 2014; Pastor *et al*., 2010; Schwibbert *et al*., 2011). They are in example implemented in microbiological and biotechnological processes for optimization of PCRs, purification of proteins, preservation of cells and for enhancement of protein activities, in example during depletion of fatty acids (Cyplik *et al*., 2012; Graf *et al*., 2008; Lentzen and Schwarz, 2006). But foremost, both chemical chaperones are nowadays used in cosmetics due to their positive effect on the human dermis, for example in skin care products and suntan lotions (bitop AG, Witten, Germany). The addition of ectoines demonstrably strengthens membrane structures, stabilizes proteins and nucleic acids, increases the moisture content of the skin and consequently protects from cold, heat and UV-light damaging (Buenger and Driller, 2004; Ablinger *et al*., 2012). Due to these properties medical applications of these two compounds as natural moisturizers and bio-stabilizers are also actively pursued, for instance in special ointments against neurodermatitis, anti-inflammatory inhalational sprays for the mouth- and nasopharyngeal zone (Sydlik *et al*., 2009) and as ingredients in nasal sprays (Olynth Ectomed). A variety of further medical application areas is currently proven and clinically tested (Freimark *et al*., 2011; Harishchandra *et al*., 2010; Sydlik *et al*., 2009; Wei *et al*., 2009). For example, research connected with Alzheimer's disease found

ectoines to effectively inhibit aggregation and neuro toxicity of prions, the major cause of this degenerative disease (Kanapathipillai *et al*., 2008; Kanapathipillai *et al*., 2005). The industrial demand in ectoines is mainly satisfied by the bitop AG company (Witten, Germany), which produces both solutes on an industrial scale (tons) through a biotechnological process called ‘bacterial milking’. In this procedure the moderate halophile *Halomonas elongata* is grown in high salinity medium (15-20% w/v NaCl) resulting in natural production and intracellular accumulation of both compounds. An artificially introduced down-shock of the high-salinity grown bacteria then leads to extrusion of the intracellular accumulated ectoine-hydroxyectoine mixture via mechanosensitive channels. While ectoine and hydroxyectoine are subsequently purified, filtered, evaporated and finally crystallized from the solute-mixture, the biomass is recycled by repatriation into the fermenter and used for a new fermentation process (Lentzen and Schwarz, 2006; Sauer and Galinski, 1998).

XIII. Aims of this dissertation

Previous studies have shown that ectoine and hydroxyectoine are highly potent stress protectants used by microorganisms to fend off the detrimental effects caused by environmental stresses (Bursy *et al*., 2008; Bursy *et al*., 2007; Kuhlmann and Bremer, 2002; Kuhlmann *et al*., 2008a; Kuhlmann *et al*., 2011; Peters *et al*., 1990). Due to their many potential industrial applications, there is a rising microbiological and biotechnological interest in understanding the molecular details and biochemical characteristics of their biosynthetic enzymes. Since the EctA and EctB enzymes possess counterparts in various metabolic pathways, the ectoine synthase EctC and the ectoine hydroxylase EctD can be considered as the key enzymes in ectoine and hydroxyectoine biosynthesis, respectively. Initial biochemical characterization of both, the ectoine synthase and the ectoine hydroxylase, has already been carried out (Bursy *et al*., 2007; Pittelkow, 2011; Moritz *et al*., 2015; Ono *et al*., 1999; Reshetnikov *et al*., 2011; Reshetnikov *et al*., 2006). In addition, crystallization trials of both enzymes have been performed as well, however, yet no structure with bound ligands has been solved (Reuter *et al*., 2010). Hence, progress has been made with respect to the biochemical characterization of these key enzymes in ectoine biosynthesis, but a full understanding of their phylogenetic distribution, main enzyme activity and structure is still lacking.

Since a deeper understanding of the key enzymes in ectoine biosynthesis is desirable, both with respect to basic science and industrial implications, the aims of this dissertation were: i) assessment of the phylogenetic distribution of ectoine and hydroxyectoine biosynthetic genes among microorganisms, ii) determination of the biochemical and kinetic properties of EctC and EctD, iii) investigations on structure-functional relationships of the key enzymes (EctC and EctD) in ectoine biosynthesis, and iv) providing a basis for the solution of the crystal structures of the ectoine synthase and the ectoine hydroxylase.

To get insights on the distribution of ectoine biosynthesis among microorganisms, phylogenetic analysis of the ectoine and hydroxyectoine biosynthetic enzymes were performed by bioinformatic approaches. Based on these results, site-directed mutagenesis of highly conserved amino acid residues of EctD were conducted to identify amino acids that play an important role in ligand-binding and are crucial for catalysis and enzyme function. Furthermore, the biochemical properties and kinetic parameters of representatives of EctC and EctD were determined. To get an overview and an in-depth understanding of their biochemistry, enzymes were chosen that on the one hand derived from microorganisms thriving habitats with extreme environmental conditions and on the other hand were widely distributed based on the results obtained by the bioinformatic analysis. Finally, identified robust enzymes were used for further crystallization trials in order to solve crystal structures (in collaboration with Dr. Sander Smits, University of Düsseldorf) of the ectoine synthase and the ectoine hydroxylase.

Results/Publications

I. Publication 1

The first publication is entitled ‘Molecular dynamics simulations and structure-guided mutagenesis provide insight into the architecture of the catalytic core of the ectoine hydroxylase’. It was published in the ‘Journal of Molecular Biology’ (doi: 10.1016/j.jmb.2013.10.028).

Widderich, N., Pittelkow, M., Hoeppner, A., Mulnaes, D., Buckel, W., Gohlke, H., Smits, S.H., and Bremer, E. Molecular dynamics simulations and structure-guided mutagenesis provide insight into the architecture of the catalytic core of the ectoine hydroxylase. *J Mol Biol*. (2014). 426: 586-600. © Elsevier.

This paper is based on a crystal structure of the ectoine hydroxylase from *Virgibacillus salexigens* that has prior been solved by Reuter and colleagues (Reuter *et al.*, 2010) and which revealed the positioning of the catalytically important iron co-factor within its active site but lacked both the substrate ectoine and the co-substrate 2-oxoglutarate. In this publication structural comparison, molecular dynamics simulations and site-directed mutagenesis experiments were used, based on the reported EctD structure, to assess the likely coordination of the substrates within the active site of the ectoine hydroxylase. This study provided a first view into the catalytic core of EctD and suggested an intricate network of interactions that coordinates the stereo-specific hydroxylation of ectoine. Based on these findings, a proposal for the EctD-catalyzed enzyme reaction was made.

Personal contribution:
I performed site-directed mutagenesis experiments (together with Marco Pittelkow), conducted production and purification of the EctD enzyme and its mutant derivatives, determined their metal content and assessed their kinetic properties by enzyme assays (together with Marco Pittelkow). I also conducted the database searches and analyzed the experimental data (together with the other authors). Figures and Tables were prepared by me (and Sander Smits). Furthermore, I contributed to the writing of the manuscript together with Sander Smits and my PhD supervisor Erhard Bremer.

Article

Molecular Dynamics Simulations and Structure-Guided Mutagenesis Provide Insight into the Architecture of the Catalytic Core of the Ectoine Hydroxylase

Nils Widderich[1,2,3,†], **Marco Pittelkow**[1,†], **Astrid Höppner**[4], **Daniel Mulnaes**[5], **Wolfgang Buckel**[1,2,6], **Holger Gohlke**[5], **Sander H.J. Smits**[7] and **Erhard Bremer**[1,2]

*1 - **Department of Biology,** Laboratory for Microbiology, Philipps-University Marburg, 35037 Marburg, Germany*
*2 - **LOEWE Center for Synthetic Microbiology,** Philipps-University Marburg, 35037 Marburg, Germany*
*3 - **Max Planck Institute for Terrestrial Microbiology Emeritus Group R. K. Thauer,** 35043 Marburg, Germany*
*4 - **X-ray Facility and Crystal Farm,** Heinrich-Heine-University Düsseldorf, 40225 Düsseldorf, Germany*
*5 - **Institute for Pharmaceutical and Medicinal Chemistry,** Heinrich-Heine-University Düsseldorf, 40225 Düsseldorf, Germany*
*6 - **Max Planck Institute for Terrestrial Microbiology,** 35043 Marburg, Germany*
*7 - **Institute of Biochemistry,** Heinrich-Heine-University Düsseldorf, 40225 Düsseldorf, Germany*

Correspondence to Erhard Bremer: *bremer@staff.uni-marburg.de*
http://dx.doi.org/10.1016/j.jmb.2013.10.028
Edited by D. Case

Abstract

Many bacteria amass compatible solutes to fend-off the detrimental effects of high osmolarity on cellular physiology and water content. These solutes also function as stabilizers of macromolecules, a property for which they are referred to as chemical chaperones. The tetrahydropyrimidine ectoine is such a compatible solute and is widely synthesized by members of the *Bacteria*. Many ectoine producers also synthesize the stress protectant 5-hydroxyectoine from the precursor ectoine, a process that is catalyzed by the ectoine hydroxylase (EctD). The EctD enzyme is a member of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily. A crystal structure of the EctD protein from the moderate halophile *Virgibacillus salexigens* has previously been reported and revealed the coordination of the iron catalyst, but it lacked the substrate ectoine and the co-substrate 2-oxoglutarate. Here we used this crystal structure as a template to assess the likely positioning of the ectoine and 2-oxoglutarate ligands within the active site by structural comparison, molecular dynamics simulations, and site-directed mutagenesis. Collectively, these approaches suggest the positioning of the iron, ectoine, and 2-oxoglutarate ligands in close proximity to each other and with a spatial orientation that will allow the region-selective and stereo-specific hydroxylation of (4*S*)-ectoine to (4*S*,5*S*)-5-hydroxyectoine. Our study thus provides a view into the catalytic core of the ectoine hydroxylase and suggests an intricate network of interactions between the three ligands and evolutionarily highly conserved residues in members of the EctD protein family.

Introduction

The exposure of cell-walled microorganisms to high osmolarity surroundings poses a considerable challenge to the physiology of the bacterial cell since it invariably elicits water efflux. The resulting loss of turgor impairs growth and threatens viability [1,2]. Many microorganisms offset these detrimental effects by amassing, either through synthesis or through uptake, compatible solutes, a class of organic osmolytes with special physiochemical properties [3–5]. Compatible solutes do not only serve as water-attracting osmolytes [1,2] but they are also highly compliant with cellular physiology and can promote and preserve the functionality of individual proteins [6–8], of protein complexes [9], and even of entire cells [10], a property for which the term "*chemical chaperone*" has been coined in the literature [11,12].

Ectoine [(4*S*)-2-methyl-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid] is an important representative of

the compatible solutes [3,13]. Its properties as a superb stabilizer of macromolecules [14] and excellent cytoprotectant [15–17] led to efforts to produce it biotechnologically on an industrial scale using highly salt tolerant microorganisms [15,18] and fostered commercial applications (e.g., in skin-care products, as PCR enhancers, or as protein and cell stabilizers); the use of ectoine for medical purposes is also currently being pursued [15–17]. Ectoine is synthesized almost exclusively by members of the *Bacteria*, and they produce it, or import it, as a protectant against stress caused by high salinity/osmolarity [16,18–25] or to relieve cellular challenges imposed by extremes in high or low growth temperatures [19,26–28]. Ectoine is synthesized from the precursor L-aspartate-β-semialdehyde, a central intermediate of microbial amino acid metabolism [29,30], by three successive steps that are catalyzed by the EctB, EctA, and EctC enzymes [23,31].

Ectoine can be classified either as a heterocyclic amino acid or as a partially hydrogenated pyrimidine derivative [22]. A substantial subgroup of the ectoine producers also synthesize a hydroxylated derivative of ectoine, 5-hydroxyectoine [(4*S*,5*S*)-5-hydroxy-2-methyl-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid] [32], through a region-selective and stereo-specific enzymatic reaction from ectoine [20,33]. Despite their close relatedness in chemical structure, 5-hydroxyectoine often possesses stabilizing and stress-protecting properties superior to those of ectoine [19,34–36]. Efforts in several laboratories led to the identification of the structural gene (*ectD*) for the ectoine hydroxylase (EctD) in various microorganisms [20,27,37], and database searches for EctD-related proteins showed that this enzyme is evolutionarily well conserved [20,38].

The biochemical purification and enzymatic characterization of the EctD protein from the moderate halophile *Virgibacillus salexigens* [20] and from the soil bacterium *Streptomyces coelicolor* [19] identified this enzyme as a member of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily (EC 1.14.11). Members of this superfamily are versatile biocatalysts whose enzyme reactions include hydroxylations, dimethylations, desaturations, epimerizations, cyclizations, halogenations, and ring expansions [39–41]. Despite these diverse biochemical activities, the so far determined crystal structures of members of this family are all closely related, and common principles of their enzyme reaction mechanism have emerged [40,42,43]. In this vein, EctD follows the enzyme reaction scheme (e.g., for the taurine dioxygenase TauD [44,45]) of this class of enzymes in catalyzing the O_2-dependent hydroxylation of ectoine. This reaction relies on a mononuclear iron catalyst and uses 2-oxoglutarate as a co-substrate; EctD thereby forms CO_2, succinate, and 5-hydroxyectoine (Fig. 1) [20,38].

The crystal structure of the biochemically characterized EctD from *V. salexigens* (formally taxonomically classified as *Salibacillus salexigens*) [20] (PDB accession code: 3EMR) is the only available structure of an ectoine hydroxylase [38]. It revealed a protein fold for the monomeric EctD [20,38] that is commonly observed among non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases [39–41,43]. The EctD crystal structure [38] consists of a double-stranded β-helix (DSBH) core, also referred to as the jelly roll or cupin fold [40], which is decorated with and stabilized by a number of α-helices. The DSBH in EctD is formed by four-stranded antiparallel β-sheets and is arranged in form of a β-sandwich (Fig. 2a).

The EctD structure also revealed the unambiguous positioning of the iron ligand within the active site, but it lacked the substrate ectoine and the co-substrate 2-oxoglutarate [38]. Within the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily, the catalytically important iron ligand is typically coordinated by side chains from a conserved HxD/E…H motif, the so-called 2-His-1-carboxylate facial triad [39,40,42,43]. The HxD/E…H motif is also present in the EctD protein [20], and the iron co-factor is coordinated by the functional groups of His146, Asp148, and His248 (Fig. 2a) that protrude into the ligand-binding cavity (Fig. 2b). The iron-coordinating residues His146 and Asp148 are part of a strictly conserved signature sequence motif for EctD-type hydroxylases of 17 amino acids in length (extending from Phe143 to Pro159) [20,38]. This consensus sequence spans an extended α-helix and a linked short β-strand within the EctD crystal structure (Fig. 2a), thereby lining one site of the presumed active site of the ectoine hydroxylase [38].

Since a deeper understanding of EctD is desirable, both with respect to basic science and industrial applications, we used the crystal structure of the *V. salexigens* EctD protein as a template for molecular docking experiments, molecular dynamics (MD) simulations, and structure-guided site-directed mutagenesis to provide deeper insight into the architecture of the catalytic core of the EctD enzyme. Emerging from these studies is a model that allows the positioning of the substrate ectoine and the co-substrate 2-oxoglutarate within the active site of the ectoine hydroxylase in such a way that it will permit the region-selective and stereo-specific hydroxylation of (4*S*)-ectoine to (4*S*,5*S*)-5-hydroxyectoine.

Results and Discussion

A view of the protein space of EctD-type enzymes

We used the *V. salexigens* EctD protein (accession number: AY935522) [20] as the query sequence

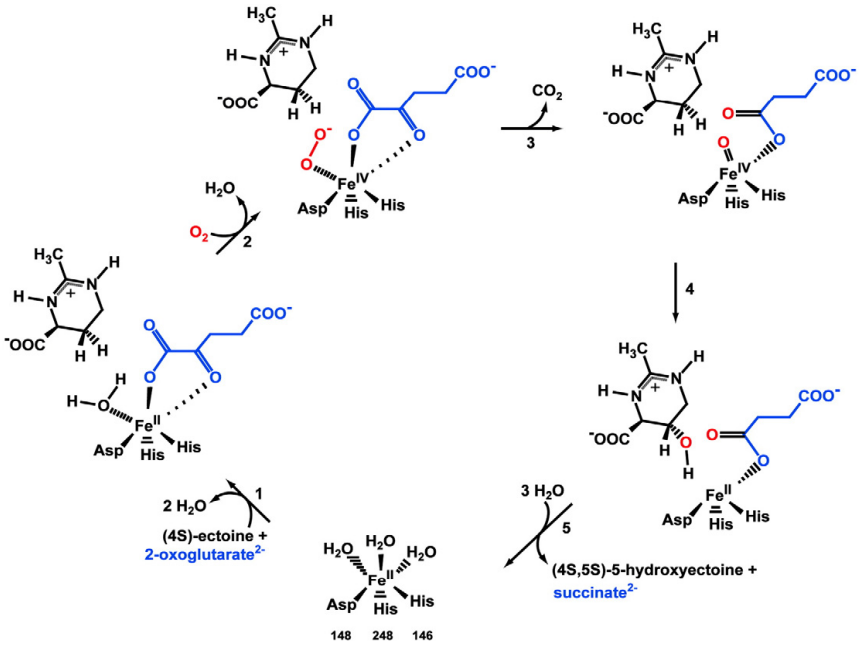


Fig. 1. EctD-catalyzed hydroxylation of ectoine. In members of non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily, a mononuclear ferrous iron center catalyzes the oxidative decarboxylation of the co-substrate 2-oxoglutarate, and this reaction is coupled with a two-electron oxidation of the substrate [40,43]. The co-substrate 2-oxoglutarate and the substrate appear to bind in a sequential fashion leading to the formation of a ternary enzyme-2-oxoglutare-substrate complex to which molecular oxygen can then bind. The proposal for the EctD-catalyzed hydroxylation of ectoine shown here is adapted from the well-established enzyme reaction mechanisms of the taurine dioxygenase TauD [44,45]: (1) The substrate-free EctD enzyme contains Fe^{II}, 6-fold-coordinated by the 2-His-1-carboxylate facial triad formed by the side chains of His146, Asp148, and His248, as well as three water molecules [38]. Two of the water molecules are replaced from the iron center upon addition of 2-oxoglutarate. Subsequently, the substrate ectoine is envisioned to bind to the EctD enzyme. (2) Upon substrate binding, dioxygen replaces the third water molecule from the iron center and oxidizes Fe^{II} to Fe^{IV}. (3) The formed Fe^{IV}-peroxo species attacks with its anion the carbonyl-C of 2-oxoglutarate to yield succinate, CO_2, and the ferryl species (Fe^{IV} = O) [46]. The ferryl species then hydroxylates ectoine to (4*S*,5*S*)-5-hydroxyectoine. (4) This may occur either by abstraction of the *pro-S* hydrogen at C5 by Fe^{IV} = O and re-addition of the hydroxyl radical from Fe^{III}-OH to the ectoine radical intermediate with retention of configuration (re-bound mechanism) or by insertion of the ferryl oxygen atom into the C–H bond (two-state reactivity mechanism) [47]. (5) Finally, the product (4*S*,5*S*)-5-hydroxyectoine and the co-product succinate are released. Three water molecules will occupy then again the empty coordination site of Fe^{II}.

in a BLAST search of the Joint Genome Institute database of microorganisms with a finished genome sequence. This search retrieved 184 proteins related to the *V. salexigens* EctD protein. An alignment of these amino acid sequences (Supplementary Fig. S1) with the ClustalW algorithm [48] showed that they possess a degree of sequence identity that ranges from 71% for the EctD protein from *Bacillus pseudofirmus* OF4 to 49% for the EctD protein from *Planctomyces maris* (DSM8797).

Since 5-hydroxyectoine is synthesized directly from ectoine [20], each of the microorganisms that possess an *ectD* gene also possess *ectC*, the structural gene for the ectoine synthase, the diagnostic enzyme for the ectoine biosynthetic route [18,21,23,24,31,49,50]. This strict correlation between the presence of the EctD and EctC proteins was ascertained by using the EctC amino acid sequence of *V. salexigens* [20] as the search query (data not shown). In each of the retrieved EctD-related proteins, the previously identified signature sequence (F*X*WHSDFETWH*X*EDGM/LP) for ectoine hydroxylases [20,38] is fully conserved. The *V. salexigens* EctD protein consists of 300 amino acid residues, and this is approximately the typical size of EctD-type proteins (Supplementary Fig. S1). Among the 185 aligned amino acid sequences, 29 amino acid

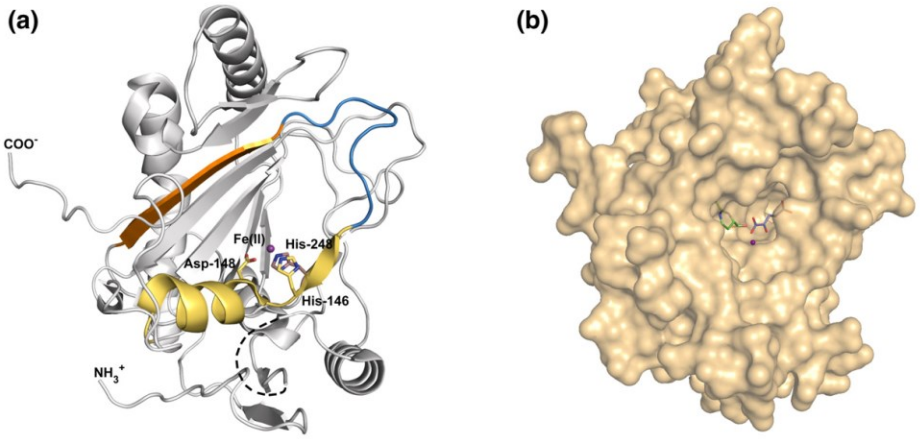


Fig. 2. Overall structure of the ectoine hydroxylase EctD from the moderate halophile *V. salexigens*. (a) The crystal structure of the EctD protein (PDB accession code: 3EMR) [38] is shown in a ribbon presentation format. The region of EctD serving as the signature sequence of ectoine hydroxylases [20,38] (highlighted in yellow) and its connection (marked in blue) to a β-sheet (highlighted in orange) containing residues important for ectoine and 2-oxoglutarate binding are emphasized. The iron(II) species is represented as a purple sphere. The region indicated by the broken line represents a flexible loop (Gly195–Leu211) that is not resolved in the EctD crystal structure [38]. (b) Surface representation of the EctD protein containing the iron(II) ligand [38] and the modeled 2-oxoglutarate and ectoine ligands.

residues are strictly conserved, including 14 residues that are part of the ectoine hydroxylase signature sequence [20,38]. Taken together, the inspection of the 185 aligned amino acid sequences revealed a strong evolutionary conservation of EctD-type proteins and a strict correlation of their occurrence with the presence of the ectoine biosynthetic enzymes in the inspected microbial genomes. Thus, we suggest that the 184 proteins identified in our database analysis are all *bona fide* ectoine hydroxylases. Each of these proteins bears the characteristic residues of enzymes that belong to the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily [40–43].

The iron-binding site

To probe the functional role of His146, Asp148, and His248 in the *V. salexigens* EctD protein for iron binding and enzyme activity [20], we separately replaced each of these amino acids by an Ala residue. None of these EctD variants was catalytically active (Table 1) nor did any of the mutant proteins contain iron (data not shown). In contrast, recombinant preparations of the wild-type *V. salexigens* EctD enzyme isolated from *Escherichia coli* cells by affinity chromatography on a Strep-Tactin resin typically had an iron content up to 0.9 mol per mole of protein [20].

In some members of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily, a Glu residue can functionally substitute for the Asp residue forming part of the iron-binding site [40,42,43]. Thus, we replaced Asp148 by a Glu residue; however, this EctD variant did not show enzyme activity either (Table 1). Collectively, we conclude from these data that no variations in the side chains forming the iron(II)-binding center of the EctD enzyme are tolerated and that all three residues forming the 2-His-1-carboxylate facial triad in the EctD protein [38] are required for iron binding and enzyme activity.

The 2-oxoglutarate-binding site

Non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases crystallized in complex with either 2-oxoglutarate or analogs possess a similarly structured binding site for this co-substrate [40,43]. Within this binding site, the 2-oxo carboxylate interacts with the iron atom in a bidentate manner; the C-5 carboxylate group points toward the inside of the active-site pocket and is bound via electrostatic and hydrogen bonding interactions. Based on the information gained from available crystal structures of members of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases [40,43], the spatial positioning of 2-oxoglutarate relative to the iron ligand in the EctD active site [38] should be predictable with a high degree of confidence.

Table 1. Kinetic properties of investigated EctD mutant derivatives.

Mutant	K_m (mM ectoine)	V_{max} (U mg^{-1})	Proposed role in the binding of the following ligands
Wild type	5.9 ± 0.3	6.2 ± 0.2	—
H146A	—	—	Fe(II)
D148A	—	—	Fe(II)
D148E	—	—	Fe(II)
H248A	—	—	Fe(II)
F95A	—	—	2-Oxoglutarate
R131A	—	—	2-Oxoglutarate
N133A	—	—	2-Oxoglutarate
F143A	—	—	2-Oxoglutarate
F143Y	—	—	2-Oxoglutarate
F143W	—	—	2-Oxoglutarate
S250A	—	—	2-Oxoglutarate
R259A	—	—	2-Oxoglutarate
R259K	—	—	2-Oxoglutarate
R259H	—	—	2-Oxoglutarate
R259Q	—	—	2-Oxoglutarate
Q129A	—	—	Ectoine
W152A	—	—	Ectoine
W152F	—	—	Ectoine
W152Y	—	—	Ectoine
S165A	17.3 ± 1.0	2.3 ± 0.6	Ectoine
F242A	—	—	Ectoine
F242Y	19.6 ± 0.5	1.7 ± 0.3	Ectoine
F242W	—	—	Ectoine
F263A	—	—	Ectoine
F263Y	8.0 ± 0.2	3.7 ± 0.4	Ectoine
F263W	—	—	Ectoine
P198A	7.1 ± 0.7	4.8 ± 0.4	Loop region
S205A	6.8 ± 0.2	5.4 ± 0.2	Loop region
A163C	6.0 ± 0.1	5.9 ± 0.4	Not involved in ligand binding
S167A	6.9 ± 0.2	6.0 ± 0.1	Not involved in ligand binding
S244C	6.2 ± 0.4	6.0 ± 0.5	Not involved in ligand binding
N261A	5.5 ± 0.4	6.4 ± 0.3	Not involved in ligand binding
V265A	16.2 ± 0.5	3.4 ± 0.2	Not involved in ligand binding
V265T	9.7 ± 0.3	5.6 ± 0.4	Not involved in ligand binding
V265L	24.6 ± 0.9	1.6 ± 0.2	Not involved in ligand binding
A163C, S244C	4.5 ± 0.6	8.1 ± 0.8	Not involved in ligand binding

The enzymatic activity of the wild-type EctD enzyme and its mutant derivatives was assayed as described in Materials and Methods. The kinetic parameters were determined by systematically varying the substrate (ectoine) concentration between 0 and 40 mM. EctD mutants that showed no activity under these conditions were considered to be catalytically inactive (–).

A search via the Dali server [51] for crystal structures related to EctD revealed the human peroxisomal phytanoyl-CoA 2-hydroxylase PhyH [52] (also known as PAXH) [53] and its relative PhyHD1A [54] as the most closely related structures to EctD with *Z*-scores of 22.6 and 22.0, respectively, and RMSD values of 2.5 Å (over 190 C^{α} atoms). While the structural relatedness of the human PhyH and PhyHD1A to EctD is substantial, the degree of their amino acid sequence identity is moderate (19% and 24%, respectively). Both the PhyH and PhyHD1A proteins have been crystallized in complex with 2-oxoglutarate [52,54], and we used the PhyHD1A crystal structure (PDB accession code: 2OPW) [54] as a point of reference to gain information about the potential binding site of 2-oxoglutarate within the EctD enzyme.

The side chains of Arg257, Ser248, and Trp174 of PhyHD1A position 2-oxoglutarate in the vicinity of the catalytically important iron ligand (Fig. 3a). The corresponding residues in EctD are Arg259, Ser250, and Phe143 (Fig. 3b). We separately replaced these residues in EctD by an Ala residue and found that each of these three EctD mutant proteins was catalytically inactive (Table 1). The 2-oxoglutarate ligand present in the PhyHD1A crystal structure contacts directly, via a charge-assisted hydrogen bond, the side chain of Trp174 (Fig. 3a). Such an interaction is not found in the predicted 2-oxoglutarate-binding site of EctD; instead, Phe143 located on the opposite side of the 2-oxoglutarate molecule makes a stacking interaction with this ligand via its aromatic side chain (Fig. 3b). We replaced Phe143 with the aromatic residues Tyr or Trp, but these EctD variants were enzymatically inactive (Table 1). Hence, Phe143 seems to play a specific role for the activity of the ectoine hydroxylase. This suggestion is supported by the complete conservation of this residue among 185 inspected EctD-type proteins (Supplementary Fig. S1). Taking the structural comparisons and the site-directed mutagenesis data together, we suggest that Arg259, Ser250, and Phe143 constitute the 2-oxoglutarate-binding site of the ectoine hydroxylase (Fig. 3b).

Besides the interactions outlined above, the 2-oxoglutate molecule also interacts with Arg131 of the EctD protein. The exact positioning of the Arg131 side chain within the ligand-binding cavity is stabilized through a cation–π interaction with the side chain of Phe95, as well as an interaction with the side chain of Asn133 (Fig. 3b). This latter interaction ensures the exact positioning of the Arg131 side chain. The lack of the Asn133 side chain results in the loss of activity (Table 1) likely due to the higher flexibility of Arg131. As to the former interaction, whereas Arg131 is located within the DSBH core of the EctD protein, Phe95 is positioned in a β-sheet flanking this fold. Hence, the Arg131/Phe95 interaction probably stabilizes the conformation of the EctD protein as a whole. In turn, loss of this interaction will likely lay the active site of the EctD protein more open and will thereby probably render it unable to bind any of its substrates. Consistent with this suggestion, Ala substitution mutations of Arg131 or Phe95 led to enzymatically inactive EctD variants (Table 1). We are thus tempted to speculate that Arg131 senses the presence of the 2-oxoglutarate

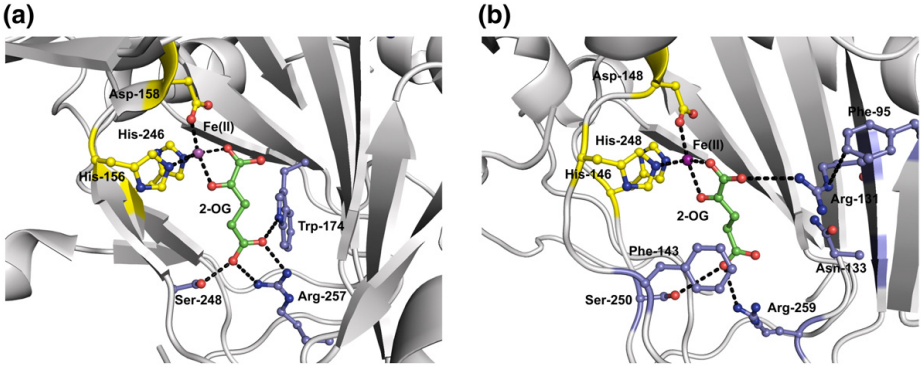


Fig. 3. Comparison of the iron(II) and 2-oxoglutarate-binding site of EctD with that present in PhyHD1A. (a) Crystal structure of a human phytanoyl-CoA dioxygenase (PhyHD1A) (PDB accession code: 3OBZ) with highlighted residues involved in the binding of iron(II) (purple sphere) and the co-substrate 2-oxoglutarate (2-OG) [54]. (b) Model of the EctD ectoine hydroxylase highlighting those residues predicted to be involved in 2-oxoglutarate (2-OG) and iron (purple sphere) binding.

co-substrate and ensures, via its interaction with Phe95, a more compact (or stable) structure of the EctD protein that is then primed for the hydroxylation of ectoine.

Inspection of the amino acid sequences of the aligned 185 EctD-type proteins (Supplementary Fig. S1) revealed that, besides the amino acid residues involved in iron binding (His146, Asp148, and His248), the amino acids predicted to be involved in 2-oxoglutarate binding (Arg259, Ser250, Phe143, Arg131, Asn133, and Phe95) (Fig. 3b) are also completely conserved (with one exception at position Ser250).

MD simulations and structure-guided mutagenesis to locate the ectoine-binding site

One of the hallmarks of compatible solutes is their preferential exclusion from the immediate hydration shell of proteins due to unfavorable interactions between these compounds and the protein backbone [6,9]. Despite the preferential exclusion of compatible solutes from protein surfaces, high-affinity interactions between these compounds and defined ligand-binding sites do occur as demonstrated by bacteria acquiring these compounds through efficient transport processes [1,3,28,55,56]. Four crystal structures for ectoine and 5-hydroxyectoine ligand-binding proteins that are associated with either bacterial ABC transporters (EhuB and OpuCC) [57,58] or TRAP transporters (UehA and TeaA) [59,60] have been reported. Although the architecture of the ectoine/5-hydroxyectoine-binding sites in these proteins varies somewhat [59], common principles for ligand binding have emerged. The carboxylate moiety of ectoine or 5-hydroxyectoine is typically coordinated within the ligand-binding site through a salt bridge with a positively charged residue. Additional salt bridges between the amidino moieties of the ectoine or 5-hydroxyectoine ligands and the substrate-binding protein provide an additional level of stabilization. A key factor contributing to ligand binding are cation–π interactions [61] that are formed between the delocalized positive charges of the ectoine and 5-hydroxyectoine molecules and aromatic side-chain residues of the ligand-binding protein. Furthermore, hydrophobic interactions and hydrogen bonds support ligand binding [57–60].

In order to generate suggestions for the location of the ectoine-binding site in EctD, we employed MD simulations. This technique has recently been successfully applied to identify ligand-binding sites within a Src kinase [62] and an alanopine dehydrogenase [63].

To generate an EctD:ectoine starting structure for the MD simulations, we first modeled the loop region between Gly195 and Leu211 of the EctD protein as these residues are missing in the available crystal structure of the *V. salexigens* ectoine hydroxylase (see Materials and Methods for details) [38]. Disordered regions corresponding to this missing loop in EctD also occur in the ligand-free crystal structures of the human phytanoyl-CoA 2-hydroxylase PhyH [52], the asparagine hydroxylase AsnO from *S. coelicolor* [64], and the AsnO-related VioC L-arginine oxygenase from *Streptomyces vinaceus* [65]. These flexible regions became ordered and structurally resolved once the PhyH, AsnO, and VioC proteins were crystallized with their corresponding ligands. The loop regions in these enzymes might thus serve as a lid to enclose the active site and shield the reactants from the surrounding

solvent during enzyme catalysis [43,52,64,65]. After careful relaxation of this EctD model, we then placed the ectoine molecule manually at a site (Supplementary Fig. S2) that we deemed suitable for ligand binding based upon principles revealed through the structural analysis of the EhuB, OpuCC, UehA, and TeaA ectoine/hydroxyectoine-binding proteins [57–60].

The *in silico* constructed EctD:ectoine complex structure (Supplementary Fig. S2) was then subjected to three independent MD simulations of 300 ns length each. These simulations revealed minor structural changes in the protein part when one does not consider the modeled loop region and the C-terminus of EctD, with RMSD values of the coordinates of C^{α} atoms (RMSD) of almost always <2.0 Å (Supplementary Fig. S3a and Fig. S4a and b). The C-terminus of the EctD protein showed a higher mobility with an RMSD up to 6 Å (data not shown); even larger movements are observed in the modeled loop region (RMSD up to 18 Å; Supplementary Fig. S3c). Although the loop moves toward the binding site in the course of the MD simulations, in none of the three trajectories did we observe that the loop actually closed the binding site as a lid, nor were interactions between loop residues and the ectoine ligand observed.

In the MD simulations, the ectoine ligand initially relocates from its manually chosen starting position (Supplementary Fig. S2) to a nearby sub-pocket formed by the side chains of residues Gln129, Trp152, Ser165, Phe242, and Phe263 characterized by RMSD values between 5 and 6 Å (Fig. 4a). It occasionally moves back toward the starting position (RMSD ≈ 3 Å) after which it returns again to the above-mentioned sub-pocket (Fig. 4a). The movement of ectoine to the sub-pocket was observed in all three independent trajectories (Fig. 4a and Supplementary Fig. S4a and b) and thereby indicated that the ectoine molecule was highly mobile, particularly at the end of simulation 1 (Fig. 4a) and around 75 and 225 ns in simulation 3 (Supplementary Fig. S4b). This high mobility of the ectoine ligand suggests a low binding affinity of the EctD protein for its substrate; this is consistent with a high K_m value, which is between 2.6 and 6 mM, depending on the particular assay conditions used and the enzyme studied [19,20]. We note that, despite the high mobility of the ectoine molecule in the MD simulations, the ligand repeatedly returns to the above-mentioned sub-pocket (as indicated by RMSD values between 5 and 6 Å; Fig. 4a and Supplementary Fig. S4b), suggesting that this sub-pocket is the preferred location for ectoine binding by the EctD protein (Fig. 4d).

Within this presumed ligand-binding site (Fig. 4d), ectoine preferentially orients such that its methyl group points to the backbone region of Ser130 at the bottom of the sub-pocket [with a distance of <5 Å between the ligands methyl group and C^{α} of Ser130 observed in ~50% of all conformations (Fig. 4b)]; its carboxylate group forms hydrogen bonds with the side chain of Ser165 in ~15% of all conformations [considering a cutoff distance of 3.2 Å between the donor and acceptor atoms (Fig. 4c)]. This binding mode leads to infrequent hydrogen bonds between the amidino moiety of ectoine and the side chain of Gln129 (~2% occurrence; data not shown). Interactions of the carboxylate group with other potential hydrogen bond donors in the binding site region, in particular, Arg131 and Ser167, are negligible (below 0.1% occurrence in all three trajectories). In this orientation, the *pro-S* hydrogen at C5 of the tetrahydropyrimidine ring of ectoine is oriented toward the catalytic iron atom (Fig. 4d), fully consistent with the fact that this hydrogen of the ectoine molecule gets replaced by a hydroxyl group during the EctD-catalyzed reaction cycle leading to the formation of (4*S*,5*S*)-5-hydroxyectoine [20,33].

To substantiate the ectoine-binding site model derived from the MD simulations (Fig. 4d), we carried out a series of site-directed mutagenesis experiments and tested the resulting EctD variants in enzyme activity assays (Table 1). Direct interactions are predicted to occur between the ectoine substrate and the side chains of Gln129 and Ser165; in addition, Trp152, Phe242, and Phe263 form the wall of the above-mentioned sub-pocket (Fig. 4d). The individual substitution of Gln129, Trp152, Phe242, and Phe263 with an Ala residue resulted in a complete loss of enzyme activity, and the replacement of Ser165 with an Ala residue substantially impaired the functioning of the EctD protein (Table 1). The conservative substitution of Trp152 with either Phe or Tyr residues resulted in enzymatically inactive EctD mutant proteins (Table 1). This can be readily understood in view of our model because the second ring of the Trp152 side chain is not present in the Phe or Tyr residues but it is this second ring that is in contact with the ectoine molecule (Fig. 4d).

Consistent with the suggestion of a functional role of Phe242 and Phe263 for ectoine binding (Fig. 4d), their individual replacement with Ala residues abolished EctD enzyme activity (see above; Table 1). We also exchanged both Phe242 and Phe263 with either Tyr or Trp residues. The Phe242/Tyr mutant is about 3-fold less active than the wild-type enzyme, whereas the Phe263/Tyr variant possesses wild-type enzymatic activity (Table 1). Conversely, the substitution of either Phe242 or Phe263 by Trp residues completely abolished the enzyme activity of the EctD mutant proteins (Table 1). Our data thus indicate that the architecture of the EctD ectoine-binding site does not tolerate larger structural modifications such as the insertion of the bulky side chain of a Trp residue. The ectoine ligand also contacts Ser130; since this contact is

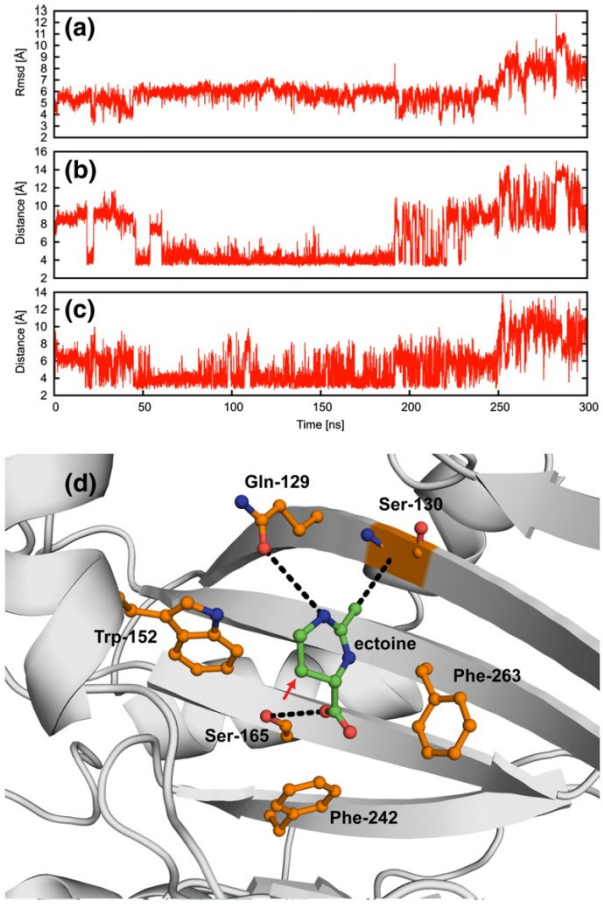


Fig. 4. MD simulation of ectoine binding by the EctD protein.(a–c) Structural parameters during the MD simulation 1 of the EctD:ectoine complex. (a) All-atom RMSD of ectoine after superimposing the core region (see Materials and Methods) of EctD on itself. (b) Distance between the methyl carbon of ectoine and the C^{α} atom of Ser130. (c) Respective minimal distance between the oxygen atoms of the carboxylate group of ectoine and the hydroxyl oxygen of Ser165. (d) Representative structure for the binding of ectoine within the EctD active site showing the dominant interactions to Ser130 and Ser165 and the occasional hydrogen bond to Gln129; in addition, Trp152, Phe242, and Phe263 are shown, which contribute to the formation of the sub-pocket containing the ectoine ligand. The read arrowhead indicates the position of the C5 atom within the ectoine molecule that becomes hydroxylated in the course of the enzyme reaction.

mediated via the protein main chain of EctD (Fig. 4d), no site-directed mutagenesis of Ser130 was carried out.

Taken together, site-directed mutagenesis of practically all residues implicated by the MD simulation in ectoine binding yield EctD variants that are either completely enzymatically inactive or strongly impaired in their function (Table 1). Furthermore, all those residues that we implicate in ectoine binding are very strongly conserved in the compiled 185 EctD-type proteins (Supplementary Fig. S1).

In our site-directed mutagenesis experiments, we specifically targeted residues that were deemed to be important for the binding of the iron, 2-oxoglutarate, and ectoine ligands and the replacement of these functionally important residues yielded almost invariably catalytically inactive EctD variants

(Table 1). While this is not surprising, it may raise the concern that essentially any amino acid substitution in the EctD protein would render the ectoine hydroxylase non-functional. We therefore carried out a series of additional mutagenesis studies of residues whose side chains either are located further away from the catalytic core of the EctD enzyme or protrude into the cupin barrel but do not make direct contacts to any of the ligands. All 10 additional mutants constructed, including residues Pro198 and Ser205 that are present within the flexible loop region of the EctD protein [38], were enzymatically active (Table 1).

Description of the architecture of the active site of the ectoine hydroxylase

The crystal structure of the EctD enzyme [38] adheres to the common protein fold of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily where a series of antiparallel β-sheets form one side of a barrel-like structure [40,43] (Fig. 1a). The side chains of those residues that we implicate in binding of the iron, 2-oxoglutarate, and ectoine ligands all protrude into the inside of the cupin barrel (Fig. 5) forming the deep ligand-binding cavity of the EctD enzyme (Fig. 2b). Key contributions to the precise positioning of these ligands are mediated by residues present on an extended α-helix and a flanking short β-strand, which contain most of the residues making up the signature sequence of EctD-type hydroxylases [20,38], and β-sheet number I, which forms the other side of the cupin barrel (Fig. 1a). The extended α-helix and its flanking short β-strand comprise residues important for iron (His146 and Asp148), 2-oxoglutarate (Phe143), and ectoine binding (Trp152) (Fig. 5). β-Sheet number I contains three strictly conserved residues (Gln129, Arg131, and Asn133) of EctD-type proteins (Supplementary Fig. S1). Gln129 is suggested to make a direct contact to the ectoine ligand (Fig. 4d), Arg131 is implicated in contacting the 2-oxoglutarate

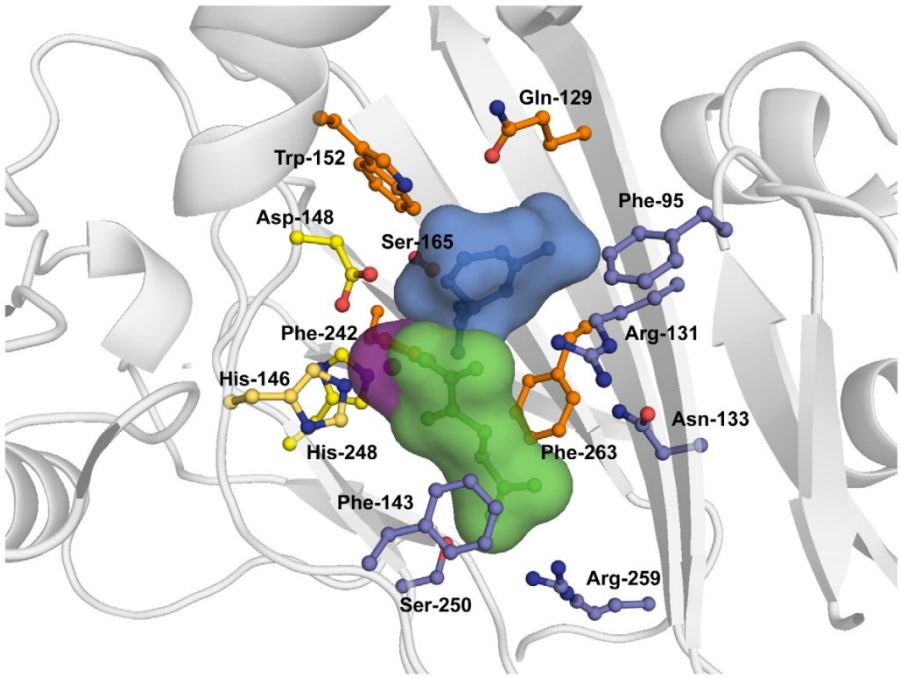


Fig. 5. Model for the catalytic core of the EctD hydroxylase. This model summarizes the predictions made by MD simulations for the positioning of the substrate ectoine (shown as blue surface), the docking of the co-substrate 2-oxoglutarate (shown as green surface), and the experimentally determined binding site for the iron(II) ligand (shown as magenta surface) [38]. The amino acids involved in binding of ectoine (orange), 2-oxoglutarate (blue), and iron (yellow) are shown in a ball-and-stick representation.

co-substrate, and Asn133 is thought to stabilize the proper conformation of the side chain of Arg131 (Fig. 3b).

Concluding remarks

We performed MD simulations starting from an *in silico* model of ectoine bound to the *V. salexigens* EctD enzyme that was inspired by the architecture of the ectoine-binding site in crystal structures of solute receptors operating in conjunction with transport systems for this compatible solute [57–60]. However, the conducted MD simulations suggest a coordination of the ectoine ligand within the active site of the ectoine hydroxylase (Fig. 4d) that is different from that observed in the aforementioned solute receptor proteins. In particular, cation–π interactions between the delocalized positive charge of the ectoine ring and aromatic residues present in the ligand-binding proteins [57–60] are not observed in the predicted ectoine-binding site of EctD (Fig. 4d). It is highly likely that these different modes of ectoine binding will contribute to the strong differences in affinity for the ectoine ligand, which are in the micromolar range for the EhuB, UehA, and TeaA solute receptors [57,59,60] and in the millimolar range for the two biochemically characterized EctD enzymes [19,20].

When one views our model for the EctD active site as a whole (Fig. 5), an intricate network of interactions between the iron, 2-oxoglutarate, and ectoine ligands and their binding partners within the catalytic core of the ectoine hydroxylase becomes apparent. This network of interactions mediates the positioning of the ectoine substrate in such a way that the C5-carbon atom in the pyrimidine ring of ectoine can be hydroxylated to (4*S*,5*S*)-5-hydroxyectoine. This hydroxylation reaction (Fig. 1) occurs with high region-selectivity and stereo-specificity both *in vivo* (e.g., in *Streptomyces parvulus*) [33] and *in vitro* by the purified *V. salexigens* EctD enzyme [20] that we have studied here. Nevertheless, the architecture of the EctD active site (Fig. 5) is probably flexible enough to allow the accommodation and hydroxylation of synthetic ectoine derivatives with reduced or expanded ring sizes [66,67] that then might possess stress-protective and protein-stabilizing properties different, and perhaps even superior, from those exhibited by their natural counterparts ectoine and 5-hydroxyectoine. The residues that we link to the binding of the iron catalyst, the 2-oxoglutarate co-substrate, and the substrate ectoine (Fig. 5) are all highly conserved among members of the EctD family of dioxygenases (Supplementary Fig. S1), suggesting that the data presented here for the EctD protein from the moderate halophile *V. salexigens* will have general functional implications for ligand binding and catalysis by ectoine hydroxylases.

Materials and Methods

Chemicals

Ectoine and 5-hydroxyectoine were kindly provided from Dr. Thomas Lentzen and Dr. Irina Bagyan (Bitop AG, Witten, Germany). 2-Oxoglutarate (disodium salt) was purchased from Sigma-Aldrich (St. Louis, MO, USA). Anhydrotetracycline hydrochloride (AHT), desthiobiotin, and Strep-Tactin Superflow chromatography material were obtained from IBA GmbH (Göttingen, Germany).

Culture conditions

E. coli strains were maintained on LB agar plates and were grown in LB liquid medium at 37 °C under aeration [68]. To select for *E. coli* DH5α (Invitrogen, Karlsruhe, Germany) cells harboring pBJ10 (*ectD*$^{+}$)-derived plasmids containing the *V. salexigens ectD* gene [20] and its mutant derivatives, we added ampicillin (100 μg ml^{-1}) to liquid and solid media. Minimal medium A [68] containing 0.5% (w/v) glucose as carbon source, 0.5% (w/v) casamino acids, 1 mM $MgSO_4$, and 3 mM thiamine was used for overproduction of the *V. salexigens* EctD protein and its mutant variants.

Site-directed mutagenesis of the *ectD* gene

To identify amino acids that might be involved in binding of the different substrates of the ectoine hydroxylase, we performed site-directed mutagenesis of the cloned *ectD* gene from *V. salexigens* [20] using the QuikChange Lightning site-directed mutagenesis kit (Stratagene, La Jolla, CA) and custom-synthesized mutagenic primers (typical length: 25 nucleotides) purchased from Biomers (Ulm, Germany). Plasmid pBJ10 [20] was used as the DNA template for the mutagenesis procedure. To ascertain the presence of the desired mutation and the absence of unwanted alterations in the coding region of the recombinant *ectD* gene, we determined the DNA sequence of the entire coding region of each mutant (Eurofins MWG GmbH, Ebersberg, Germany). The following *ectD* variants were generated in the parent plasmid pBJ10: pMP1 (Phe^{143} → Ala [TTT → GCG]), pMP2 (Arg^{259} → Ala [CGT → GCG]), pMP3 (Phe^{143} → Tyr [TTT → TAC]), pMP4 (Arg^{259} → Lys [CGT → AAG]), pMP5 (Arg^{259} → His [CGT → CAC]), pMP6 (Arg^{259} → Gln [CGT → CAA]), pMP7 (Phe^{263} → Ala [TTC → GCC]), pMP8 (Asn^{261} → Ala [AAT → GCT]), pMP9 (Asn^{133} → Ala [AAT → GCT]), pMP10 (Arg^{131} → Ala [CGA → GCA]), pMP11 (Trp^{152} → Ala [TGG → GCG]), pMP12 (Gln^{129} → Ala [CAA → GCA]), pMP13 (Ser^{167} → Ala [TCT → GCT]), pMP14 (Phe^{95} → Ala [TTT → GCG]), pMP15 (Phe^{143} → Trp [TTT → TGG]), pMP17 (His^{146} → Ala [CAT → GCT]), pMP18 (Asp^{148} → Glu [GAT → GAG]), pMP19 (Asp^{148} → Ala [GAT → GCT]), pMP20 (His^{248} → Ala [CAT → GCT]), pMP21 (Phe^{263} → Tyr [TTC → TAC]), pMP22 (Phe^{263} → Trp [TTC → TGG]), pMP23 (Trp^{152} → Phe [TGG → TTC]), pMP24 (Trp^{152} → Tyr [TGG → TAC]), pMP26 (Ser^{205} → Ala [TCA → GCA]), pMP29 (Pro^{189} → Ala [CCT → GCT]), pMP30 (Ser^{250} → Ala [TCA → GCA]), pWN1 (Phe^{242} → Ala [TTT → GCT]), pWN2 (Phe^{242} → Tyr [TTT → TAT]) pWN3 (Phe^{242} → Trp [TTT → TGG]). pWN4 (Ala^{163} → Cys

[GCC → TGC]), pWN5 (Ser^{244} → Cys [AGT → TGT]), pWN6 (Ala^{163} → Cys [GCC → TGC]; Ser^{244} → Cys [AGT → TGT]), pWN16 (Val^{265} → Ala [GTG → GCG]), pWN17 (Val^{265} → Thr [GTG → ACG]), pWN18 (Val^{265} → Leu [GTG → CTG]), and pWN19 (Ser^{165} → Ala [AGT → GCT]).

Overproduction and purification of recombinant EctD proteins

Plasmid pBJ10 is a derivative of the pASK-IBA3 expression vector (IBA) and harbors the wild-type *ectD* gene from the moderate halophile *V. salexigens* (DSM 11483^{T}) [20]. In this plasmid, transcription of the otherwise osmotically inducible *ectD* gene [20] is mediated by the anhydrotetracycline (AHT)-inducible *tet*-promoter present on the pASK-IBA3 expression vector; hence, the production of the EctD protein can be triggered by adding the inducer (AHT) of the TetR repressor protein to the culture medium of *E. coli* strains. To permit the purification of the EctD protein by affinity chromatography, we fused the *ectD* coding region present on plasmid pBJ10 at its 3′ end to a short DNA sequence encoding a *Strep*-tag-II peptide. Overexpression of the recombinant *ectD* gene and its mutant derivatives was performed in the host *E. coli* strain DH5α essentially as described by Reuter *et al.* [38]. Typically, 1 l of minimal medium A was inoculated to an OD_{578} of 0.1 from an overnight culture of DH5α (pBJ10 or its mutant derivatives) propagated in the same medium. When the culture (grown at 37 °C) had reached mid-log phase (OD_{578} of about 0.7), the high-level expression of *ectD* was triggered by the addition of AHT (final concentration of 0.2 mg ml^{-1}), the growth temperature was reduced to 35 °C, and the cells were then grown for additional 2 h. The cells were harvested by centrifugation (10 min, 5000 rpm, 4 °C), and the cell pellets were stored at −20 °C until further use. Purification of EctD was attained by affinity chromatography on a Strep-Tactin Superflow column as detailed previously [20]. Similar yields of the wild-type and mutant EctD proteins were obtained by this purification scheme, indicating that the single amino mutations introduced into EctD did not grossly affect protein stability. The purified EctD proteins were shock-frozen in liquid nitrogen and stored at −80 °C until their enzyme activity was determined via an HPLC-based assay [20,26].

EctD enzyme assays

The enzymatic activity of EctD and its mutant derivatives was assayed in 30 µl reaction volumes that typically contained 10 mM TES buffer (pH 7.5), 10 mM 2-oxoglutarate, 6 mM ectoine, 1 mM $FeSO_4$, and various amounts (between 2 and 15 µg) of the recombinantly produced EctD protein. In contrast to previously used assay conditions [20], the reactive oxygen species scavenger catalase was left out. To determine the kinetic parameters of the *V. salexigens* EctD wild-type enzyme and its mutant derivatives, we varied the concentration of ectoine in the assay between 0 and 40 mM. The enzyme reaction mixtures were incubated at 32 °C for 20 min in a thermomixer (Eppendorf, Hamburg, Germany) with vigorous shaking to provide enough oxygen to the O_2-dependent EctD enzyme. The enzyme reactions were stopped by adding 30 µl of 100% acetonitrile to the reaction mixtures, and the samples were immediately centrifuged to remove the denatured EctD protein (10 min, 4 °C, 32,000*g*). Formation of 5-hydroxyectoine from the substrate ectoine was determined by loading 20 µl of the supernatant onto a GROM-SIL 100 Amino-1PR column (125 mm × 4 mm, 3-µm particle size; GROM, Rottenburg-Hailfingen, Germany) attached to a UV-visible detector system (LINEAR UVIS 205; SYKAM, Fürstenfeldbruck, Germany). Ectoine and 5-hydroxyectoine were monitored by their absorbance at 210 nm [20,21] and quantitated with the ChromStar 7.0 software package (SYKAM). The iron content of EctD protein preparations produced in *E. coli* by recombinant techniques was assayed as previously described [69].

Modeling of the flexible loop region (Gly195–Leu211) of the EctD protein

In order to generate a starting structure for MD simulations of the ectoine ligand bound to the EctD enzyme of *V. salexigens*, we took the coordinates of the EctD protein atoms from the PDB entry 3EMR [38]. In the reported EctD structure, residues Gly195 to Leu211 have not been resolved, which indicates the presence of a mobile loop region [38]. In order to generate coordinates of this region by homology modeling, we searched suitable template structures in the PDB by means of a BLAST search. This resulted in PDB entries 2WBO (9% sequence identity according to the Dali Structural Alignment Server), 2WBP (10%), and 2WBQ (10%), that is, crystal structures of the non-heme iron(II) oxygenase VioC from *S. vinaceus* involved in the synthesis of the antibiotic viomycin [65]. As may be expected from the low sequence identity of the EctD and VioC proteins, we did not succeed in generating a multiple sequence alignment of the template and target sequences that resulted in meaningful models, neither by way of automatic alignments with ClustalW or Modeller nor by way of a structural alignment using the Dali Server, nor by combining automatic and manual alignment. Thus, we resorted to the modeling of the missing loop residues of the EctD protein with the automatic loop refinement of Modeller [70]. Initially, Modeller's automodel class was used with a 3-fold slow refinement for 300 iterations with a molpdf ("energy" computed by Modeller's objective function) limit of 106; subsequently, the loops in each model were refined using the dope_loopmodel [71] class with the slow md_level, giving 25 models in total.

Each model was processed using Anolea [72], Dope [73], and PROCHECK [74] to determine structure quality. Anolea provides a residue-wise score; Dope, a residue-wise score and a global score. For PROCHECK a residue-wise score was calculated as the maximum deviation of dihedral angles, bond angles, and bond lengths from optimal values for that residue. The residue-wise scores are smoothed over a window of 9 residues, and "Ensemble *Z*-scores" are computed by comparing these scores for each residue in a model to corresponding residues in all models; averaging over all residues yield the global "Ensemble Scores". Finally, a "Composite Ensemble Score" is computed as the average of the Anolea, Dope, and PROCHECK scores. The "best" model was defined as the one with the lowest "Composite Ensemble Score". This model is characterized by an average Anolea energy per residue of −10.6, a global Dope score of −1.8903,

and a Ramachandran plot with 90.8% of the residues in the core region, 8.8% in the allowed region, 0.0% in the generously allowed region, and 0.4% in the disallowed region. This model was then used for further generation of the *in silico* EctD/ectoine complex structure. In this model, the loop is not occluding the binding site region.

MD simulations

MD simulations were performed with the AMBER 11 suite of programs [75] using the GPU accelerated code of *pmemd* [76]. For the EctD protein, the force field by Cornell *et al.* [77] was used with modifications suggested by Simmerling *et al.* [78]. For ectoine and 2-oxoglutarate, the general amber force field was used [79]; partial charges were generated by the RESP procedure [80]. The catalytic metal ion was modeled with parameters for Mg^{2+}. The two His residues coordinating the metal ion were protonated at $N^{\delta 1}$ (His248) and $N^{\varepsilon 2}$ (His146); all other His residues were protonated at $N^{\varepsilon 2}$. Visual inspection confirmed that the metal ion remained in the binding site throughout the simulations. The structures were placed into an octahedral periodic box of TIP3P water molecules [81] such that the distance between the edges of the water box and the closest atom of the EctD protein was at least 11 Å. This resulted in a system size of ~40,000 atoms. The particle mesh Ewald method [82] was used to treat long-range electrostatic interactions. Bond lengths involving bonds to hydrogen atoms were constrained using SHAKE [83]. The timestep for all MD simulations was 2 fs, with a direct-space, non-bonded cutoff of 8 Å.

In order to relax the modeled loop, we energy minimized the *apo* structure of EctD by 500 steps of conjugate gradient minimization, applying harmonic restraints with force constants of at least 5 kcal mol^{-1} $Å^{-2}$ to all solute atoms. Applying harmonic restraints with force constants of 5 kcal mol^{-1} $Å^{-2}$ to all solute atoms, we carried out NVT-MD for 50 ps, during which the system was heated from 100 K to 300 K. Subsequent NPT-MD was used for 150 ps to adjust the solvent density. Finally, the force constants of the harmonic restraints on solute atom positions were gradually reduced to 1 kcal mol^{-1} $Å^{-2}$ during 50 ps of NVT-MD, followed by 10 ns of NVT-MD with all atoms, but the modeled loop restrained in that way, again followed by 200 ns of NVT-MD without any restraints.

The structure obtained at the end of this procedure was aligned to the PDB entry 3EMR. The coordinates of the catalytic metal ion were copied, and ectoine was then placed in the presumed active-site region of the EctD protein in a way that its carboxylate group forms a salt bridge with residue Arg131 and that its methyl group points toward residue Phe263. This orientation was chosen because the ectoine ligand assumes a similar configuration with the binding sites of the high-affinity ligand-binding proteins EhuB (PDB accession code: 2Q88) [57], UehA (PDB accession code: 3FXB) [59], and TeaA (PDB accession code: 2VPN) [60]. Finally, 2-oxoglutarate was placed in the structure using the program COOT [84]; this co-substrate remained largely in its predicted binding site throughout the simulations (RMSD mostly around 3 Å; Supplementary Fig. S3b). This structure was again thermalized following the procedure described above for relaxing the loop region except that now the force constants of the harmonic restraints on solute atom positions were gradually reduced to zero during 100 ps of NVT-MD.

At this point, three NVT-MD simulations for production were spawned, at 300.0, 300.1, and 300.2 K, respectively. Trajectories of 300 ns length were separately analyzed with the program *ptraj*, with conformations extracted every 20 ps. The core region of the protein was defined to be all residues but those of the loop and the mobile C-terminus.

Database searches and alignment of EctD amino acid sequences

Proteins that are homologous to the EctD protein from *V. salexigens* were searched for via the Web server of the DOE Joint Genome Institute[‡]. The amino acid sequences of the retrieved EctD-type proteins were aligned using ClustalW [48].

Preparation of figures

Figures of the experimentally determined structure of the *V. salexigens* EctD protein (PDB accession code: 3EMR) [38] or of conformations extracted from the MD trajectories were prepared using the PyMOL software package[§]. Graphs were prepared with gnuplot[‖].

Acknowledgements

We thank Jochen Sohn for excellent technical assistance and Georg Lentzen and Irina Bagyan (Bitop AG) for their kind gifts of ectoine and 5-hydroxyectoine. We greatly appreciate the expert help of Vickie Koogle in the language editing of our manuscript and thank Lutz Schmitt for helpful discussions. N.W. and E.B. are very grateful to Rolf Thauer for his continued support. E.B. greatly valued the hospitality and kind support of Tom Silhavy during a sabbatical at the Department of Molecular Biology of Princeton University (Princeton, NJ, USA). We thank the anonymous reviewers of our manuscript for their thoughtful advice. Financial support for this study was generously provided by grants from the Deutsche Forschungsgemeinschaft through the SFB 987 (to E.B.), by the LOEWE program of the state of Hessen (via the Center for Synthetic Microbiology, Marburg) (to E.B. and W.B.), the Max Planck Institute for Terrestrial Microbiology (to W.B. and N.W.), by the Fonds der Chemischen Industrie (to E.B.), by the initiative "Fit for Excellence" of the Heinrich-Heine-University of Düsseldorf (to H.G., A.H., and S.H.J.S.), and by the Ministry of Innovation, Science, and Research of North Rhine-Westphalia and Heinrich-Heine-University, Düsseldorf, for a scholarship to D.M. within the CLIB-Graduate Cluster Industrial Biotechnology. Computational support for the work of H.G. was provided by the "Zentrum für Informations und

Medientechnologie" at the Heinrich-Heine-University. N.W. is a member of the International Max Planck Research School for Environmental, Cellular and Molecular Microbiology (Marburg) and gratefully acknowledges its support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jmb.2013.10.028.

Received 30 April 2013;
Received in revised form 29 September 2013;
Accepted 23 October 2013
Available online 30 October 2013

Keywords:
dioxygenase;
compatible solute;
osmotic stress;
molecular modeling

† N.W. and M.P. contributed equally to this work.

‡ http://www.jgi.doe.gov/
§ http://www.pymol.org
‖ http://www.gnuplot.info

Abbreviations used:
DSBH, double-stranded β-helix; MD, molecular dynamics.

References

[1] Bremer E, Krämer R. Coping with osmotic challenges: osmoregulation through accumulation and release of compatible solutes. In: Storz G, Hengge-Aronis R, editors. Bacterial stress responses. Washington: ASM Press; 2000. p. 79–97.
[2] Wood JM. Bacterial osmoregulation: a paradigm for the study of cellular homeostasis. Annu Rev Microbiol 2011;65:215–38.
[3] Kempf B, Bremer E. Uptake and synthesis of compatible solutes as microbial stress responses to high osmolality environments. Arch Microbiol 1998;170:319–30.
[4] Wood JM, Bremer E, Csonka LN, Krämer R, Poolman B, van der Heide T, et al. Osmosensing and osmoregulatory compatible solute accumulation by bacteria. Comp Biochem Physiol Part A Mol Integr Physiol 2001;130:437–60.
[5] Yancey PH. Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. J Exp Biol 2005;208:2819–30.
[6] Street TO, Bolen DW, Rose GD. A molecular mechanism for osmolyte-induced protein stability. Proc Natl Acad Sci U S A 2006;103:13997–4002.
[7] Ignatova Z, Gierasch LM. Inhibition of protein aggregation *in vitro* and *in vivo* by a natural osmoprotectant. Proc Natl Acad Sci U S A 2006;103:13357–61.
[8] Bourot S, Sire O, Trautwetter A, Touze T, Wu LF, Blanco C, et al. Glycine betaine-assisted protein folding in a *lysA* mutant of *Escherichia coli*. J Biol Chem 2000;275:1050–6.
[9] Street TO, Krukenberg KA, Rosgen J, Bolen DW, Agard DA. Osmolyte-induced conformational changes in the Hsp90 molecular chaperone. Protein Sci 2010;19:57–65.
[10] Manzanera M, Garcia de Castro A, Tondervik A, Rayner-Brandes M, Strom AR, Tunnacliffe A. Hydroxyectoine is superior to trehalose for anhydrobiotic engineering of *Pseudomonas putida* KT2440. Appl Environ Microbiol 2002;68:4328–33.
[11] Diamant S, Eliahu N, Rosenthal D, Goloubinoff P. Chemical chaperones regulate molecular chaperones *in vitro* and in cells under combined salt and heat stresses. J Biol Chem 2001;276:39586–91.
[12] Chattopadhyay MK, Kern R, Mistou MY, Dandekar AM, Uratsu SL, Richarme G. The chemical chaperone proline relieves the thermosensitivity of a *dnaK* deletion mutant at 42 degrees C. J Bacteriol 2004;186:8149–52.
[13] da Costa MS, Santos H, Galinski EA. An overview of the role and diversity of compatible solutes in Bacteria and Archaea. Adv Biochem Eng Biotechnol 1998;61:117–53.
[14] Lippert K, Galinski EA. Enzyme stabilization by ectoine-type compatible solutes: protection against heating, freezing and drying. Appl Microbiol Biotechnol 1992;37:61–5.
[15] Lentzen G, Schwarz T. Extremolytes: natural compounds from extremophiles for versatile applications. Appl Microbiol Biotechnol 2006;72:623–34.
[16] Pastor JM, Salvador M, Argandona M, Bernal V, Reina-Bueno M, Csonka LN, et al. Ectoines in cell stress protection: uses and biotechnological production. Biotechnol Adv 2010;28:782–801.
[17] Graf R, Anzali S, Buenger J, Pfluecker F, Driller H. The multifunctional role of ectoine as a natural cell protectant. Clin Dermatol 2008;26:326–33.
[18] Schwibbert K, Marin-Sanguino A, Bagyan I, Heidrich G, Lentzen G, Seitz H, et al. A blueprint of ectoine metabolism from the genome of the industrial producer *Halomonas elongata* DSM 2581^T. Environ Microbiol 2011;13:1973–94.
[19] Bursy J, Kuhlmann AU, Pittelkow M, Hartmann H, Jebbar M, Pierik AJ, et al. Synthesis and uptake of the compatible solutes ectoine and 5-hydroxyectoine by *Streptomyces coelicolor* A3(2) in response to salt and heat stresses. Appl Environ Microbiol 2008;74:7286–96.
[20] Bursy J, Pierik AJ, Pica N, Bremer E. Osmotically induced synthesis of the compatible solute hydroxyectoine is mediated by an evolutionarily conserved ectoine hydroxylase. J Biol Chem 2007;282:31147–55.
[21] Kuhlmann AU, Bremer E. Osmotically regulated synthesis of the compatible solute ectoine in *Bacillus pasteurii* and related *Bacillus* spp. Appl Environ Microbiol 2002;68:772–83.
[22] Galinski EA, Pfeiffer HP, Trüper HG. 1,4,5,6-Tetrahydro-2-methyl-4-pyrimidinecarboxylic acid. A novel cyclic amino acid from halophilic phototrophic bacteria of the genus *Ectothiorhodospira*. Eur J Biochem 1985;149:135–9.
[23] Louis P, Galinski EA. Characterization of genes for the biosynthesis of the compatible solute ectoine from *Marinococcus halophilus* and osmoregulated expression in *Escherichia coli*. Microbiology 1997;143:1141–9.
[24] Reshetnikov AS, Khmelenina VN, Mustakhimov II, Kalyuzhnaya M, Lidstrom M, Trotsenko YA. Diversity and phylogeny of the ectoine biosynthesis genes in aerobic, moderately halophilic methylotrophic bacteria. Extremophiles 2011;15:653–63.

[25] Jebbar M, von Blohn C, Bremer E. Ectoine functions as an osmoprotectant in *Bacillus subtilis* and is accumulated via the ABC-transport system OpuC. FEMS Microbiol Lett 1997;154:325–30.

[26] Kuhlmann AU, Bursy J, Gimpel S, Hoffmann T, Bremer E. Synthesis of the compatible solute ectoine in *Virgibacillus pantothenticus* is triggered by high salinity and low growth temperature. Appl Environ Microbiol 2008;74:4560–3.

[27] Garcia-Estepa R, Argandona M, Reina-Bueno M, Capote N, Iglesias-Guerra F, Nieto JJ, et al. The *ectD* gene, which is involved in the synthesis of the compatible solute hydroxyectoine, is essential for thermoprotection of the halophilic bacterium *Chromohalobacter salexigens*. J Bacteriol 2006;188:3774–84.

[28] Kuhlmann AU, Hoffmann T, Bursy J, Jebbar M, Bremer E. Ectoine and hydroxyectoine as protectants against osmotic and cold stress: uptake through the SigB-controlled betaine-choline-carnitine transporter-type carrier EctT from *Virgibacillus pantothenticus*. J Bacteriol 2011;193:4699–708.

[29] Stöveken N, Pittelkow M, Sinner T, Jensen RA, Heider J, Bremer E. A specialized aspartokinase enhances the biosynthesis of the osmoprotectants ectoine and hydroxyectoine in *Pseudomonas stutzeri* A1501. J Bacteriol 2011;193:4456–68.

[30] Lo CC, Bonner CA, Xie G, D'Souza M, Jensen RA. Cohesion group approach for evolutionary analysis of aspartokinase, an enzyme that feeds a branched network of many biochemical pathways. Microbiol Mol Biol Rev 2009;73:594–651.

[31] Ono H, Sawada K, Khunajakr N, Tao T, Yamamoto M, Hiramoto M, et al. Characterization of biosynthetic enzymes for ectoine as a compatible solute in a moderately halophilic eubacterium, *Halomonas elongata*. J Bacteriol 1999;181:91–9.

[32] Inbar L, Lapidot A. The structure and biosynthesis of new tetrahydropyrimidine derivatives in actinomycin D producer *Streptomyces parvulus*. Use of ^{13}C- and ^{15}N-labeled L-glutamate and ^{13}C and ^{15}N NMR spectroscopy. J Biol Chem 1988;263:16014–22.

[33] Inbar L, Frolow F, Lapidot A. The conformation of new tetrahydropyrimidine derivatives in solution and in the crystal. Eur J Biochem 1993;214:897–906.

[34] Borges N, Ramos A, Raven ND, Sharp RJ, Santos H. Comparative study of the thermostabilizing properties of mannosylglycerate and other compatible solutes on model enzymes. Extremophiles 2002;6:209–16.

[35] Ablinger E, Hellweger M, Leitgeb S, Zimmer A. Evaluating the effects of buffer conditions and extremolytes on thermostability of granulocyte colony-stimulating factor using high-throughput screening combined with design of experiments. Int J Pharm 2012;436:744–52.

[36] Van-Thuoc D, Hashim SO, Hatti-Kaul R, Mamo G. Ectoine-mediated protection of enzyme from the effect of pH and temperature stress: a study using *Bacillus halodurans* xylanase as a model. Appl Microbiol Biotechnol 2013;97:6271–8.

[37] Prabhu J, Schauwecker F, Grammel N, Keller U, Bernhard M. Functional expression of the ectoine hydroxylase gene (*thpD*) from *Streptomyces chrysomallus* in *Halomonas elongata*. Appl Environ Microbiol 2004;70:3130–2.

[38] Reuter K, Pittelkow M, Bursy J, Heine A, Craan T, Bremer E. Synthesis of 5-hydroxyectoine from ectoine: crystal structure of the non-heme iron(II) and 2-oxoglutarate-dependent dioxygenase EctD. PLoS One 2010;5:e10647.

[39] Hausinger RP. FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes. Crit Rev Biochem Mol Biol 2004;39:21–68.

[40] Aik W, McDonough MA, Thalhammer A, Chowdhury R, Schofield CJ. Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. Curr Opin Struct Biol 2012;22:691–700.

[41] Prescott AG, Lloyd MD. The iron(II) and 2-oxoacid-dependent dioxygenases and their role in metabolism. Nat Prod Rep 2000;17:367–83.

[42] Straganz GD, Nidetzky B. Variations of the 2-His-1-carboxylate theme in mononuclear non-heme FeII oxygenases. ChemBioChem 2006;10:1536–48.

[43] Hangasky JA, Taabazuing CY, Valliere MA, Knapp MJ. Imposing function down a (cupin)-barrel: secondary structure and metal stereochemistry in the alphaKG-dependent oxygenases. Metallomics 2013;5:287–301.

[44] Knauer SH, Hartl-Spiegelhauer O, Schwarzinger S, Hanzelmann P, Dobbek H. The Fe(II)/alpha-ketoglutarate-dependent taurine dioxygenases from *Pseudomonas putida* and *Escherichia coli* are tetramers. FEBS J 2012;279:816–31.

[45] Grzyska PK, Appelman EH, Hausinger RP, Proshlyakov DA. Insight into the mechanism of an iron dioxygenase by resolution of steps following the FeIV = HO species. Proc Natl Acad Sci U S A 2010;107:3982–7.

[46] Riggs-Gelasco PJ, Price JC, Guyer RB, Brehm JH, Barr EW, Bollinger JM, et al. EXAFS spectroscopic evidence for an Fe = O unit in the Fe(IV) intermediate observed during oxygen activation by taurine:alpha-ketoglutarate dioxygenase. J Am Chem Soc 2004;126:8108–9.

[47] Schwarz H. Chemistry with methane: concepts rather than recipes. Angew Chem Int Ed Engl 2011;50:10096–115.

[48] Thompson JD, Plewniak F, Thierry JC, Poch O. DbClustal: rapid and reliable global multiple alignments of protein sequences detected by database searches. Nucleic Acids Res 2000;28:2919–26.

[49] Saum SH, Müller V. Growth phase-dependent switch in osmolyte strategy in a moderate halophile: ectoine is a minor osmolyte but major stationary phase solute in *Halobacillus halophilus*. Environ Microbiol 2008;10:716–26.

[50] Vargas C, Argandona M, Reina-Bueno M, Rodriguez-Moya J, Fernandez-Aunion C, Nieto JJ. Unravelling the adaptation responses to osmotic and temperature stress in *Chromohalobacter salexigens*, a bacterium with broad salinity tolerance. Saline Systems 2008;4:14.

[51] Holm L, Rosenström P. Dali server: conservation mapping in 3D. Nucleic Acids Res 2010;38:W545–9.

[52] McDonough MA, Kavanagh KL, Butler D, Searls T, Oppermann U, Schofield CJ. Structure of human phytanoyl-CoA 2-hydroxylase identifies molecular mechanisms of Refsum disease. J Biol Chem 2005;280:41101–10.

[53] Schofield CJ, McDonough MA. Structural and mechanistic studies on the peroxisomal oxygenase phytanoyl-CoA 2-hydroxylase (PhyH). Biochem Soc Trans 2007;35:870–5.

[54] Zhang Z, Kochan GT, Ng SS, Kavanagh KL, Oppermann U, Schofield CJ, et al. Crystal structure of PHYHD1A, a 2OG oxygenase related to phytanoyl-CoA hydroxylase. Biochem Biophys Res Commun 2011;408:553–8.

[55] Ziegler C, Bremer E, Krämer R. The BCCT family of carriers: from physiology to crystal structure. Mol Microbiol 2010;78:13–34.

[56] Bremer E. A look into the aromatic cage. Environ Microbiol 2011;3:1–5.

[57] Hanekop N, Höing M, Sohn-Bösser L, Jebbar M, Schmitt L, Bremer E. Crystal structure of the ligand-binding protein EhuB from *Sinorhizobium meliloti* reveals substrate recognition of the compatible solutes ectoine and hydroxyectoine. J Mol Biol 2007;374:1237–50.
[58] Du Y, Shi WW, He YX, Yang YH, Zhou CZ, Chen Y. Structures of the substrate-binding protein provide insights into the multiple compatible solute binding specificities of the *Bacillus subtilis* ABC transporter OpuC. Biochem J 2011;436:283–9.
[59] Lecher J, Pittelkow M, Zobel S, Bursy J, Bonig T, Smits SH, et al. The crystal structure of UehA in complex with ectoine-A comparison with other TRAP-T binding proteins. J Mol Biol 2009;389:58–73.
[60] Kuhlmann SI, Terwisscha van Scheltinga AC, Bienert R, Kunte HJ, Ziegler C. 1.55 Å structure of the ectoine binding protein TeaA of the osmoregulated TRAP-transporter TeaABC from *Halomonas elongata*. Biochemistry 2008;47:9475–85.
[61] Dougherty DA. Cation-pi interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. Science 1996;271:163–8.
[62] Shan Y, Kim ET, Eastwood MP, Dror RO, Seeliger MA, Shaw DE. How does a drug molecule find its target binding site? J Am Chem Soc 2011;133:9181–3.
[63] Gohlke H, Hergert U, Meyer T, Mulnaes D, Grieshaber MK, Smits SH, et al. Binding region of alanopine dehydrogenase predicted by unbiased molecular dynamics simulations of ligand diffusion. J Chem Inf Model 2013;53:2493–8.
[64] Strieker M, Kopp F, Mahlert C, Essen LO, Marahiel MA. Mechanistic and structural basis of stereospecific Cbeta-hydroxylation in calcium-dependent antibiotic, a daptomycin-type lipopeptide. ACS Chem Biol 2007;2:187–96.
[65] Helmetag V, Samel SA, Thomas MG, Marahiel MA, Essen LO. Structural basis for the erythro-stereospecificity of the L-arginine oxygenase VioC in viomycin biosynthesis. FEBS J 2009;276:3669–82.
[66] Schnoor M, Voss P, Cullen P, Boking T, Galla HJ, Galinski EA, et al. Characterization of the synthetic compatible solute homoectoine as a potent PCR enhancer. Biochem Biophys Res Commun 2004;322:867–72.
[67] Witt EM, Davies NW, Galinski EA. Unexpected property of ectoine synthase and its application for synthesis of the engineered compatible solute ADPC. Appl Microbiol Biotechnol 2011;91:113–22.
[68] Miller JH. Experiments in molecular genetics. New York: Cold Spring Harbor Laboratory; 1972.
[69] Lovenberg W, Buchanan BB, Rabinowitz JC. Studies on the chemical nature of *Clostridial ferredoxin*. J Biol Chem 1963;238:3899–913.
[70] Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993;234:779–815.
[71] Fiser A, Do RK, Sali A. Modeling of loops in protein structures. Protein Sci 2000;9:1753–73.
[72] Melo F, Feytmans E. Assessing protein structures with a non-local atomic interaction energy. J Mol Biol 1998;277:1141–52.
[73] Shen MY, Sali A. Statistical potential for assessment and prediction of protein structures. Protein Sci 2006;15:2507–24.
[74] Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–91.
[75] Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, et al. The Amber biomolecular simulation programs. J Comput Chem 2005;26:1668–88.
[76] Gotz AW, Williamson MJ, Xu D, Poole D, Le Grand S, Walker RC. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized Born. J Chem Theory Comput 2012;8:1542–55.
[77] Cornell WD, Cieplak CI, Bayly IR, Gould IR, Merz KM, Ferguson DM, et al. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 1995;117:5179–97.
[78] Simmerling C, Strockbine B, Roitberg AE. All-atom structure prediction and folding simulations of a stable protein. J Am Chem Soc 2002;124:11258–9.
[79] Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J Comput Chem 2004;25:1157–74.
[80] Bayly CI, Cieplak P, Cornell WD, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for determining atom-centered charges: the RESP model. J Phys Chem 1993;97:10269–80.
[81] Jorgensen WL, Chandrasekhar J, Madura J, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–35.
[82] Darden T, York D, Pedersen L. Particle mesh Ewald: an $N\cdot\log(N)$ method for Ewald sums in large systems. J Chem Phys 1993;98:10089–92.
[83] Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical-integration of Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. J Comput Phys 1977;23:327–41.
[84] Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr Sect D Biol Crystallogr 2004;60:2126–32.

Supplemental data

Molecular dynamics simulations and structure-guided mutagenesis provide insight into the architecture of the catalytic core of the ectoine hydroxylase

Nils Widderich[1,2,3◇], Marco Pittelkow[1◇], Astrid Höppner[4], Daniel Mulnaes[5], Wolfgang Buckel[1,2,6], Holger Gohlke[5], Sander H. J. Smits[7], Erhard Bremer[1,2*]

[1]Department of Biology, Laboratory for Microbiology, Philipps-University Marburg, Marburg, Germany

[2]LOEWE-Center for Synthetic Microbiology, Philipps-University Marburg, Marburg, Germany

[3]Max Planck Institute for Terrestrial Microbiology, Emeritus Group R.K. Thauer, Marburg, Germany

[4]X-Ray Facility and Crystal Farm, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

[5]Institute for Pharmaceutical and Medicinal Chemistry, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

[6]Max Planck Institute for Terrestrial Microbiology, Marburg, Germany

[7]Institute of Biochemistry, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

Running head: Ectoine Hydroxylase

[◇]Both authors contributed equally to this study

For correspondence:

Dr. Erhard Bremer, Philipps-University Marburg, Karl-von-Frisch Str. 8, D-35043 Marburg, Germany
Phone (49)-6421-282 1529; Fax: (49)-6421-282 8979; E-mail: bremer@staff.uni-marburg.de

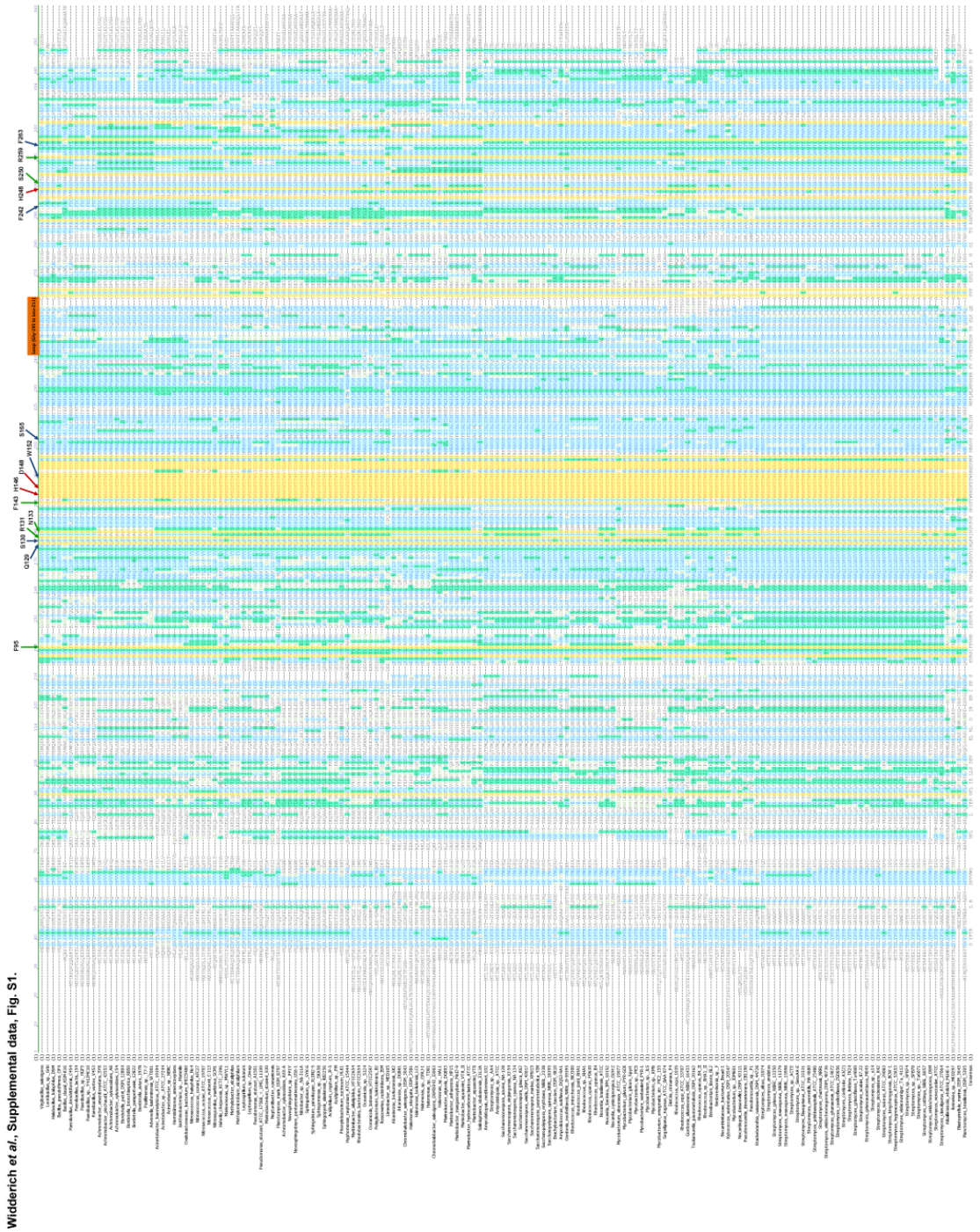


Fig. S1. Alignment of the amino acid sequences of EctD-type proteins. Yellow boxes highlight amino acids that are strictly conserved. Amino acids involved in ligand binding are indicated by a red (Fe(II)), a green (2-oxoglutarate) or a blue (ectoine) arrow. The loop region (G195-L211) is indicated by an orange box.

Widderich *et al.*, Supplemental data, Fig. S2.

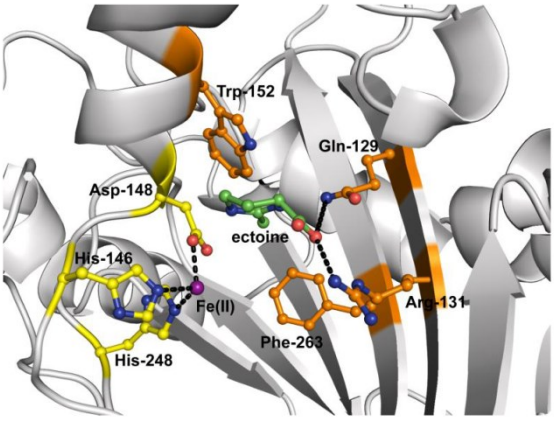


Fig. S2. Starting *in silico* EctD structure used in the MD simulations. This structure is based upon considerations gained from ectoine/hydroxyectoine binding proteins associated with transport systems. Amino acid residues possibly involved in ectoine binding are highlighted in orange.

Widderich *et al.*, Supplemental data, Fig. S3.

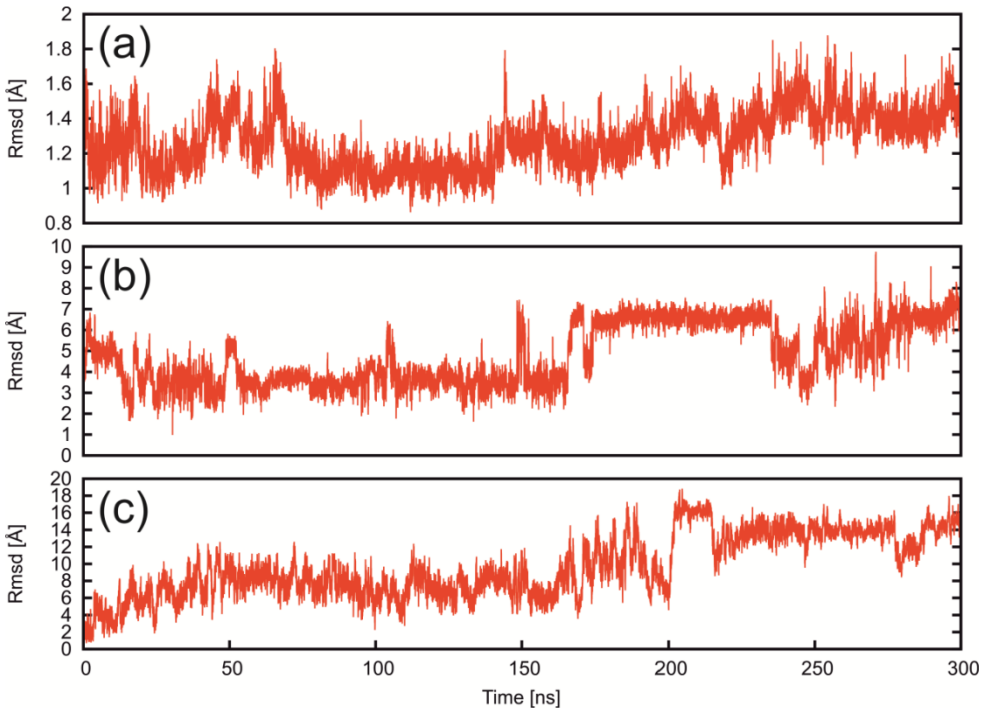


Fig. S3. Structural deviations during the MD simulation 1 of the EctD/ectoine complex. (a) RMSD (root mean square deviation) of the C_α atoms of EctD after superimposing the core region (see Materials and Methods) of EctD on itself. (b) RMSD of the 2-oxoglutarate atoms after superimposing the core region of EctD on itself. (c) RMSD of the C_α atoms of the modeled loop after superimposing the core region of EctD on itself.

Widderich *et al.*, Supplemental data, Fig. S4.

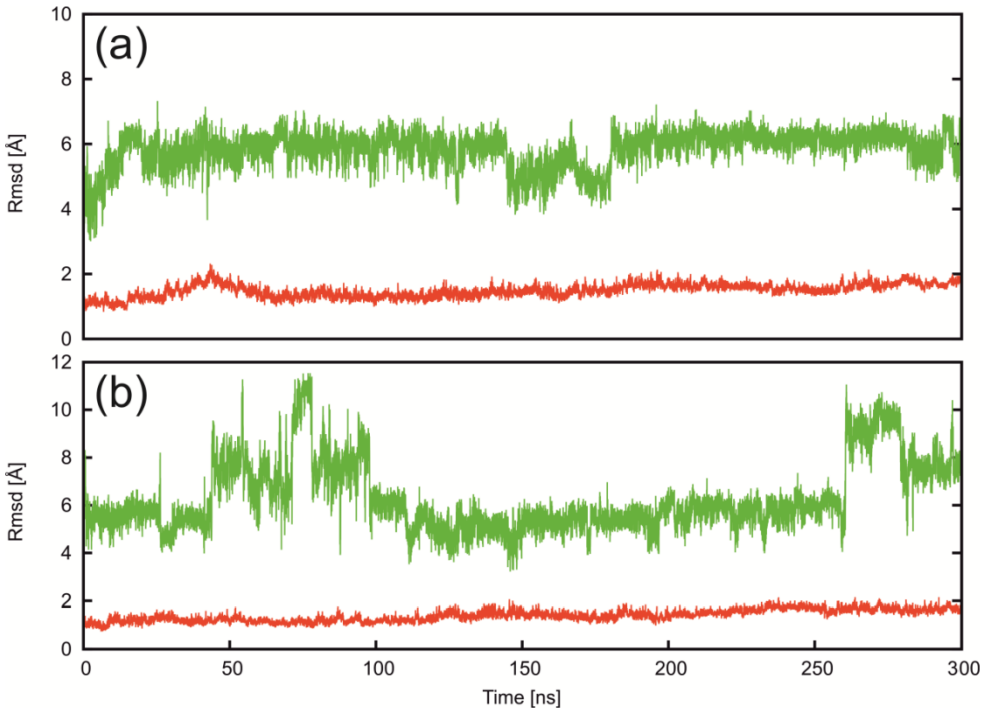


Fig. S4. Structural deviations during the MD simulations 2 & 3 of the EctD/ectoine complex. RMSD (root mean square deviation) of the C_α atoms of EctD after superimposing the core region (see Materials and Methods) of EctD on itself (red) and all-atom RMSD of ectoine after superimposing the core region of EctD on itself (green). (a) Results from MD simulation 2; (b) results from MD simulation 3.

II. Publication 2

The second publication is entitled ‘Overexpression, crystallization and preliminary X-ray crystallographic analysis of the ectoine hydroxylase from *Sphingopyxis alaskensis*’. It was published in ‘Acta Crystallographica Section F Structural Biology Communications’ (doi: 10.1107/S2053230X14004798).

Hoeppner, A., Widderich, N., Bremer, E., and Smits, S.H. Overexpression, crystallization and preliminary X-ray crystallographic analysis of the ectoine hydroxylase from *Sphingopyxis alaskensis*. *Acta Crystallogr F Struct Biol Commun*. (2014). 70: 493-496.

This paper describes a detailed expression and purification protocol for the ectoine hydroxylase from the cold-adapted bacterium *Sphingopyxis alaskensis*. It further deals with crystallization trials and preliminary X-ray analysis of the EctD enzyme resulting in two different crystal forms of the *Sa*EctD protein which represent an apo- and an iron-bound form of the protein.

Personal contribution:

I established production and purification protocols for the *Sa*EctD protein and performed the corresponding experiments as well as determined the metal content of the enzymes and analyzed the experimental data (together with the other authors). I also prepared the Figures occurring in the publication (together with Sander Smits) and contributed to the writing of the manuscript together with my PhD supervisor Erhard Bremer and our collaborator Sander Smits.

CrossMark

crystallization communications

Acta Crystallographica Section F
Structural Biology Communications
ISSN 2053-230X

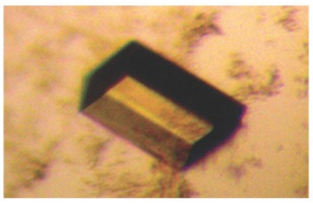
Overexpression, crystallization and preliminary X-ray crystallographic analysis of the ectoine hydroxylase from *Sphingopyxis alaskensis*

Astrid Hoeppner,[a] Nils Widderich,[b,c,d] Erhard Bremer[b,c] and Sander H. J. Smits[e]*

[a]Crystal Farm and X-ray Facility, Heinrich-Heine-Universität, Universitätsstrasse 1, 40225 Düsseldorf, Germany, [b]Department of Biology, Laboratory for Microbiology, Philipps-University Marburg, Marburg, Germany, [c]LOEWE Center for Synthetic Microbiology, Philipps-University Marburg, Marburg, Germany, [d]Max Planck Institute for Terrestrial Microbiology, Emeritus Group R. K. Thauer, Marburg, Germany, and [e]Institut für Biochemie, Heinrich-Heine-Universität, Universitätsstrasse 1, 40225 Düsseldorf, Germany

Correspondence e-mail: sander.smits@hhu.de

Received 10 January 2014
Accepted 2 March 2014



The ectoine hydroxylase (EctD) is a member of the non-haem-containing iron(II)- and 2-oxoglutarate-dependent dioxygenase superfamily. Its mononuclear iron centre is a prerequisite for the activity of this enzyme and promotes the O_2-dependent oxidative decarboxylation of 2-oxoglutarate, which is coupled to a two-electron oxidation of the substrate ectoine to yield 5-hydroxyectoine. An expression and purification protocol for the EctD enzyme from *Sphingopyxis alaskensis* was developed and the protein was crystallized using the sitting-drop vapour-diffusion method. This resulted in two different crystal forms, representing the apo and iron-bound forms of the enzyme.

1. Introduction

Ectoine and its derivative 5-hydroxyectoine are well known members of the compatible solutes and are widely produced as protectants against osmotic stress by numerous microorganisms. Ectoine is synthesized from L-aspartate β-semialdehyde by the EctABC enzymes (Louis & Galinski, 1997; Ono *et al.*, 1999). Ectoine is a superb stabilizer of macromolecules (Lippert & Galinski, 1992), an excellent cytoprotectant (Graf *et al.*, 2008; Pastor *et al.*, 2010) and is commercially used as sun protection agent in skincare products (Lentzen & Schwarz, 2006). A subgroup of ectoine producers convert ectoine to 5-hydroxyectoine, which possesses stress-protective and protein-function-preserving properties that are different from and often superior to those of ectoine (Borges *et al.*, 2002; Bursy *et al.*, 2008).

Hydroxylation of ectoine is catalyzed by the ectoine hydroxylase (EctD), a member of the non-haem-containing iron(II)- and 2-oxoglutarate-dependent dioxygenase superfamily (Prabhu *et al.*, 2004; García-Estepa *et al.*, 2006; Bursy *et al.*, 2007). This hydroxylation reaction requires O_2 and 2-oxoglutarate as co-substrates whereby CO_2, succinate and 5-hydroxyectoine are formed (Bursy *et al.*, 2007). The EctD-catalyzed reaction (Fig. 1*a*) is strictly dependent on a mononuclear iron centre promoting the O_2-dependent oxidative decarboxylation of 2-oxoglutarate, which is coupled with a two-electron oxidation of the substrate ectoine (Widderich *et al.*, 2014).

Here, we present the results of the purification, crystallization and preliminary X-ray crystallographic analyses from two different crystal types of the EctD protein from *Sphingopyxis alaskensis*, a microorganism that is well adapted to permanently cold marine environments (Ting *et al.*, 2010). To solve its crystal structure, we plan to use the method of molecular replacement with the structure of the EctD protein from *Virgibacillus salexigens* (PDB entry 3emr; Reuter *et al.*, 2010), which displays 50.8% amino-acid sequence identity to the *S. alaskensis* enzyme.

2. Materials and methods

2.1. Overexpression and purification

Plasmid pMP40 (*ectD*$^+$) was used for the overexpression of the *S. alaskensis* EctD protein (*Sa*EctD, Accession No. YP_617990). The *S. alaskensis ectD* gene was amplified *via* PCR from chromosomal DNA using custom-synthesized DNA primers (ectD_Spha_fwd, ATGGTAGGTCTCAAATGCAAGACCTCTACCCCTCGCGC;

crystallization communications

ectD_Spha_rev, ATGGTAGGTCTCAGCGCTTGCCGGCACCGTTTC-GACGAG). *Bsa*I restriction sites were synthetically added to the ends of the DNA primers, allowing cloning of the amplified full-length *ectD* gene into plasmid pASK-IBA3 and thereby fusing it to a C-terminal *Strep*-tag II affinity peptide. The resulting plasmid pMP40 carries the *S. alaskensis ectD* gene under the control of the TetR-responsive and anhydrotetracycline (AHT)-inducible *tet* promoter carried by the pASK-IBA3 plasmid backbone. As a consequence, overexpression of the cloned *S. alaskensis ectD* gene can be induced by adding AHT (purchased from IBA GmbH) to the growth medium. The *Sa*EctD-*Strep*-tag II hybrid protein was purified by affinity chromatography on Strep-Tactin Superflow material (purchased from IBA GmbH).

To provide the *Sa*EctD enzyme for crystallization trials, *Escherichia coli* BL21 (pMP40) cells were grown at 310 K in Minimal Medium A (MMA) supplemented with ampicillin (100 µg ml^{-1}) in a 2 l Erlenmeyer flask (filled with 1 l medium) in an aerial shaker set to 180 rev min^{-1}. At an OD_{578} of 0.7 of the culture, overexpression was induced by the addition of AHT (final concentration of 0.2 mg ml^{-1}); the temperature was than dropped to 303 K, the speed of the aerial shaker was reduced to 100 rev min^{-1} and the cells were propagated for an additional 2 h. The cells were then harvested by centrifugation (10 min at 4800*g* in a Hettich Rotana speed centrifuge at 277 K). The pelleted cells were resuspended in buffer *A* (20 m*M* TES pH 8, 100 m*M* KCl) and they were then disrupted by passing them three times in the cold (277 K) through a French Pressure Cell Press (SLM Aminco) at 1000 psi (1 psi = 6.895 kPa). Cellular debris was removed by ultracentrifugation (60 min at 100 000*g* and 277 K) and the cleared supernatant was loaded onto a Strep-Tactin Superflow column that had been equilibrated with five bed volumes of buffer *A*. The column was then washed with ten column volumes of buffer *A*. The EctD-*Strep*-tag II protein was eluted from the affinity chromatography material with three column volumes of buffer *A* containing 2.5 m*M* desthiobiotin. The eluted EctD-*Strep*-tag II protein was concentrated with Vivaspin 6 columns (Sartorius Stedim Biotech GmbH, Göttingen, Germany) to a concentration of about 10 mg ml^{-1} before it was used for crystallization trials. 200–300 mg EctD-*Strep*-tag II protein per litre of cell culture were routinely obtained using this overproduction and purification scheme. The protein concentration was measured using the Pierce BCA Protein Assay Kit (Thermo Scientific, Schwerte, Germany) and an extinction coefficient of 41 035 M^{-1} cm^{-1} at 280 nm and the molecular mass (35.29 kDa) of the full-length EctD including the *Strep*-tag II. The iron content of *Sa*EctD was determined as described by Lovenberg *et al.* (1963). The purity of the *Sa*EctD protein was assessed by SDS–PAGE (12% polyacrylamide; Fig. 1*b*). The *Sa*EctD protein was shock-frozen in liquid nitrogen and stored at 193 K until use for crystallization.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

	Apo *Sa*EctD	Fe-*Sa*EctD
Beamline	ID23eh2, ESRF	ID23eh2, ESRF
Detector	MAR225	MAR225
Temperature (K)	100	100
Wavelength (Å)	0.87260	0.87260
Crystal-to-detector distance (mm)	245	291
Rotation range per image (°)	0.4	0.2
Total rotation range (°)	100	120
Exposure time per image (s)	2.1	0.4
Space group	$C222_1$	$P2_12_12_1$
Unit-cell parameters		
a (Å)	83.48	78.16
b (Å)	86.51	87.52
c (Å)	95.34	96.05
$\alpha = \beta = \gamma$ (°)	90	90
Resolution (Å)	30–2.1 (2.2–2.1)	30–2.7 (2.8–2.7)
No. of observed reflections	85055	91395
No. of unique reflections	20251	18652
Mean redundancy	4.2 (4.1)	4.9 (5.0)
Completeness (%)	99.7 (99.8)	99.6 (99.9)
$\langle I/\sigma(I)\rangle$	15.1 (2.8)	19.8 (2.9)
Mosaicity (°)	0.09	0.19
R_{meas}†	6.2 (49.9)	5.8 (58.9)
R_{pim}‡	3.6 (42.0)	2.9 (47.3)
Overall B factor from Wilson plot (Å^2)	43.1	63.8
Matthews coefficient V_M (Å^3 Da^{-1})		2.41
Monomer	2.44	4.65
Dimer		2.33
Solvent content (%)	49.6	47.2

† $R_{meas} = \sum_{hkl}\{N(hkl)/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$. ‡ $R_{p.i.m.} = \sum_{hkl}\{1/[N(hkl)-1]\}^{1/2}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$.

2.2. Crystallization and preliminary X-ray analysis of *Sa*EctD

For all crystallization trials the full-length *Sa*EctD protein including the *Step*-tag II peptide was used.

2.2.1. Apo *Sa*EctD. In order to find initial crystallization conditions we used standard screening kits from Qiagen, Hilden, Germany (Nextal JCSG Core Suites I–IV) and Molecular Dimensions, Suffolk, England (MemGold, MemGold 2, MIDAS) in Corning 3553 sitting-drop plates at 293 K. 0.5 µl of the homogeneous protein solution of EctD (10 mg ml^{-1} in 20 m*M* TES pH 7.5, 80 m*M* NaCl) was mixed with 0.5 µl reservoir solution. From the roughly 700 conditions tested,

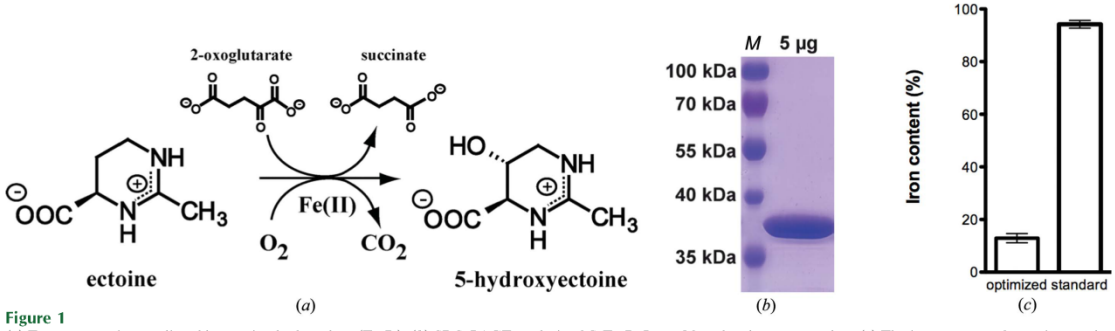


Figure 1
(*a*) Enzyme reaction mediated by ectoine hydroxylase (EctD). (*b*) SDS–PAGE analysis of *Sa*EctD. Lane *M*, molecular-mass marker. (*c*) The iron content of protein samples was measured for the optimized *Sa*EctD overproduction protocol for crystallography purposes and the standard *Sa*EctD protein production protocol used for enzyme activity assays (Widderich *et al.*, 2014).

12 conditions resulted in initial crystal hits. By optical inspection we chose four of them for optimization experiments in which we varied two parameters against each other (*e.g.* pH against PEG concentration) in sitting-drop trials at 293 K with drops composed of 1.5 µl protein solution and 1.5 µl reservoir solution. The largest and best diffracting crystals resulted from solutions consisting of 100 m*M* MES pH 6.0, 200 m*M* calcium acetate, 30%(*w*/*v*) PEG 400. To improve diffraction properties we tested this condition in combination with the additive and detergent screens from Hampton Research, Aliso Viejo, USA as described with 1/10 of the drop volume of the corresponding additive or detergent. The detergent *n*-dodecyl-*N*′,*N*-dimethylglycine yielded the crystals that diffracted the best (Fig. 2*a*). Optimized crystallization trials were then performed using the sitting-drop vapour-diffusion method at 293 K. 1.5 µl of the homogeneous protein solution of EctD (10 mg ml^{-1} in 20 m*M* TES pH 7.5, 80 m*M* NaCl) was mixed with 1.5 µl reservoir solution consisting of 100 m*M* MES pH 6.0, 200 m*M* Ca acetate, 30%(*w*/*v*) PEG 400 and 1.5 m*M* *n*-dodecyl-*N*′,*N*-dimethylglycine and equilibrated over 300 µl reservoir solution. Crystals grew within 6–12 d to their final dimensions of around 30 × 30 × 50 µm. Crystals were cryoprotected by slowly and cautiously adding 1 µl 100% glycerol to the crystallization drop by stepwise pipetting before cooling the crystals in liquid nitrogen.

2.2.2. Fe-*Sa*EctD. Starting from the optimized condition of the apo *Sa*EctD crystals, we performed crystallization trials for Fe-*Sa*EctD. The conditions were as described for the apo *Sa*EctD protein, but the protein solution was premixed with 100 m*M* $FeCl_2$ to a final concentration of 4 m*M* and incubated on ice for 10–15 min. EctD crystals were grown under the above-mentioned conditions but with 3.5 m*M* *n*-dodecyl-*N*′,*N*-dimethylglycine. They grew within 6–12 d at 293 K to final dimensions of around 40 × 40 × 180 µm (Fig. 2*b*). The crystals were cryoprotected as described above.

Data sets were collected from a single crystal of either apo *Sa*EctD or Fe-*Sa*EctD on beamline ID23eh2 at the ERSF, Grenoble, France at 100 K. These data sets were processed using the *XDS* package (Kabsch, 2010*a*) and scaled with *XSCALE* (Kabsch, 2010*b*).

3. Results

The ectoine hydroxylase from *S. alaskensis* (*Sa*EctD) was overexpressed in *E. coli* BL21 carrying plasmid pMP40 as a recombinant EctD-*Strep*-tag II protein and purified by affinity chromatography using a Strep-Tactin Superflow column. By slight variations of the initial expression protocol (Bursy *et al.*, 2007; Reuter *et al.*, 2010;

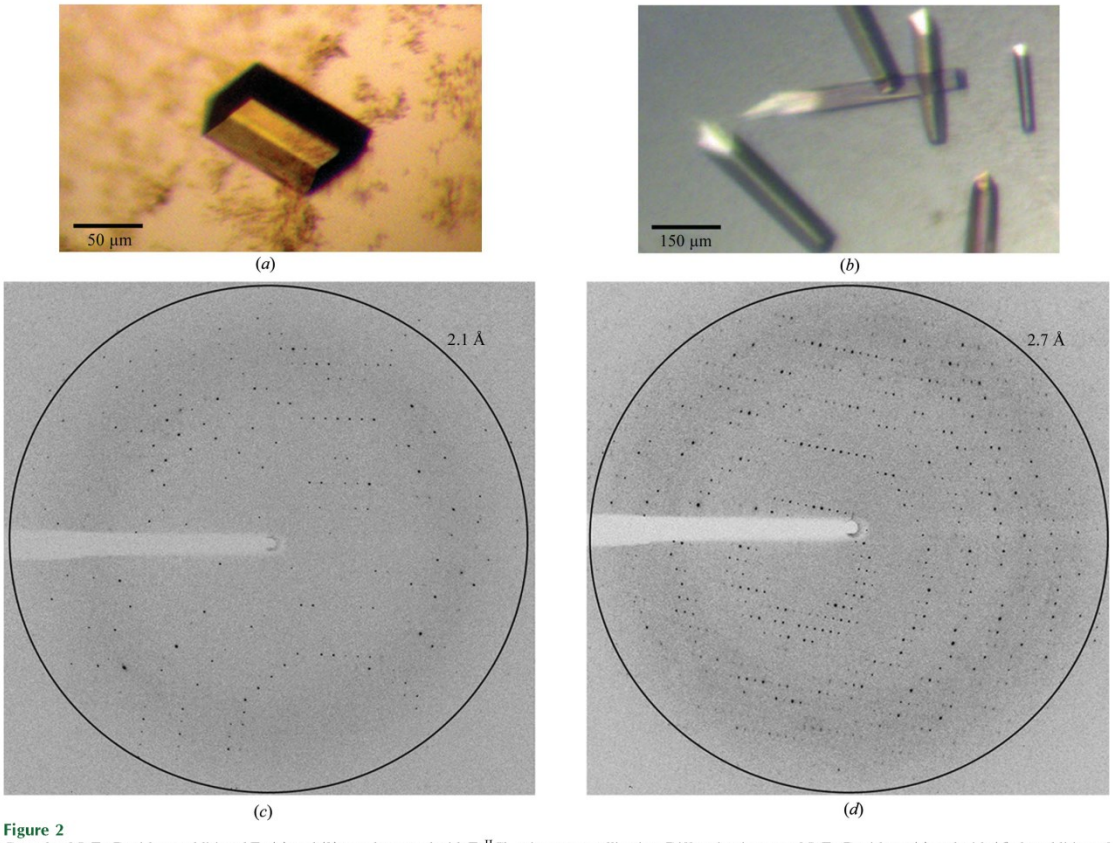


Figure 2
Crystals of *Sa*EctD without additional Fe (*a*) and (*b*) supplemented with $Fe^{II}Cl_2$ prior to crystallization. Diffraction images of *Sa*EctD without (*c*) and with (*d*) the addition of Fe ions prior to crystallization [oscillation width 0.4° (*c*) and 0.2° (*d*)].

crystallization communications

Widderich *et al.*, 2014) (see §2), the amount of purified *Sa*EctD protein was increased tenfold to 200–300 mg l^{-1}. Protein purity was validated by SDS–PAGE (Fig. 1*b*). The purity of *Sa*EctD was assumed to be at least 95% by optical inspection. The homogeneity of *Sa*EctD was analytically checked by size-exclusion chromatography.

Since the presence of a correctly complexed iron ligand is critical for EctD-mediated enzyme catalysis (Widderich *et al.*, 2014), we determined the iron content of the purified protein and found between 0.12 and 0.14 mol of iron per mole of EctD (Fig. 1*c*). This low iron content is likely to be due to the newly developed strong expression protocol, since the initial protocol (Widderich *et al.*, 2014) yielded 0.9 mol iron per mol of EctD protein (Fig. 1*c*).

Recently, it has been shown that the iron ligand can be added to the purified enzyme, leading to a restored enzyme activity of *Sa*EctD (Widderich *et al.*, 2014). For crystallization experiments we tested whether there is an effect on the crystallization behaviour and/or crystal quality when the *Sa*EctD protein was supplemented with additional Fe ions prior to crystallization.

Diffracting crystals of apo *Sa*EctD (*i.e.* no additional iron added) and Fe-*Sa*EctD (supplemented with $FeCl_2$) were obtained with 100 m*M* MES pH 6.0, 200 m*M* calcium acetate, 30%(*w*/*v*) PEG 400 and different concentrations of *n*-dodecyl-*N*′,*N*-dimethylglycine (1.5 and 3.5 m*M*) using the sitting-drop vapour-diffusion method (Figs. 2*a* and 2*b*). After adding glycerol as a cryoprotectant crystals were shock-frozen in liquid nitrogen. Native data sets for both protein crystal species were collected at 100 K. Crystals of apo *Sa*EctD diffracted to a maximum resolution of 2.1 Å and those of Fe-*Sa*EctD diffracted to a maximum resolution of 2.7 Å (Figs. 2*c* and 2*d*).

Preliminary data processing using the *XDS* package resulted in different unit-cell parameters and, more interestingly, the crystals displayed different space groups. Whereas apo *Sa*EctD crystallized in space group $C222_1$, Fe-*Sa*EctD displayed a $P2_12_12_1$ symmetry (see Table 1). The apo form contains one monomer per asymmetric unit whereas the Fe-supplemented form contains a dimer. V_M values were calculated to be 2.4 Å^3 Da^{-1} for the apo and 2.3 Å^3 Da^{-1} for the Fe-*Sa*EctD crystal with a solvent content of 50% and 47%, respectively (Matthews, 1968).

When one compares the data statistics of apo *Sa*EctD with those of Fe-treated *Sa*EctD, the slight differences in the unit-cell parameters in combination with the different space groups and the asymmetric unit content may be a hint that the apo crystals lack the iron catalyst whereas in the Fe-supplemented crystals the iron might be present. The structure determination of the *Sa*EctD protein using both crystal forms *via* molecular replacement using the *Vs*EctD structure (PDB entry 3emr; Reuter *et al.*, 2010) as a template model is currently in progress.

We thank the staff of the P14 beamline at the EMBL, Hamburg, Germany, for kind support during crystal screening. We acknowledge the European Synchrotron Radiation Facility for provision of synchrotron radiation facilities, especially Sean McSweeny from ID23eh2. We also gratefully acknowledge the 'Fit For Excellence' Fund of the Heinrich Heine University, the SFB 987, the IMPRS Marburg and the Emeritus group of R. K. Thauer for financial support. We thank Marco Pittelkow for the construction of the expression plasmid.

References

Borges, N., Ramos, A., Raven, N. D., Sharp, R. J. & Santos, H. (2002). *Extremophiles*, **6**, 209–216.
Bursy, J., Kuhlmann, A. U., Pittelkow, M., Hartmann, H., Jebbar, M., Pierik, A. J. & Bremer, E. (2008). *Appl. Environ. Microbiol.* **74**, 7286–7296.
Bursy, J., Pierik, A. J., Pica, N. & Bremer, E. (2007). *J. Biol. Chem.* **282**, 31147–31155.
García-Estepa, R., Argandoña, M., Reina-Bueno, M., Capote, N., Iglesias-Guerra, F., Nieto, J. J. & Vargas, C. (2006). *J. Bacteriol.* **188**, 3774–3784.
Graf, R., Anzali, S., Buenger, J., Pfluecker, F. & Driller, H. (2008). *Clin. Dermatol.* **26**, 326–333.
Kabsch, W. (2010*a*). *Acta Cryst.* D**66**, 125–132.
Kabsch, W. (2010*b*). *Acta Cryst.* D**66**, 133–144.
Lentzen, G. & Schwarz, T. (2006). *Appl. Microbiol. Biotechnol.* **72**, 623–634.
Lippert, K. & Galinski, E. A. (1992). *Appl. Microbiol. Biotechnol.* **37**, 61–65.
Louis, P. & Galinski, E. A. (1997). *Microbiology*, **143**, 1141–1149.
Lovenberg, W., Buchanan, B. B. & Rabinowitz, J. C. (1963). *J. Biol. Chem.* **238**, 3899–3913.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Ono, H., Sawada, K., Khunajakr, N., Tao, T., Yamamoto, M., Hiramoto, M., Shinmyo, A., Takano, M. & Murooka, Y. (1999). *J. Bacteriol.* **181**, 91–99.
Pastor, J. M., Salvador, M., Argandoña, M., Bernal, V., Reina-Bueno, M., Csonka, L. N., Iborra, J. L., Vargas, C., Nieto, J. J. & Cánovas, M. (2010). *Biotechnol. Adv.* **28**, 782–801.
Prabhu, J., Schauwecker, F., Grammel, N., Keller, U. & Bernhard, M. (2004). *Appl. Environ. Microbiol.* **70**, 3130–3132.
Reuter, K., Pittelkow, M., Bursy, J., Heine, A., Craan, T. & Bremer, E. (2010). *PLoS One*, **5**, e10647.
Ting, L., Williams, T. J., Cowley, M. J., Lauro, F. M., Guilhaus, M., Raftery, M. J. & Cavicchioli, R. (2010). *Environ. Microbiol.* **12**, 2658–2676.
Widderich, N., Pittelkow, M., Höppner, A., Mulnaes, D., Buckel, W., Gohlke, H., Smits, S. H. J. & Bremer, E. (2014). *J. Mol. Biol.* **426**, 586–600.

III. Publication 3

The third publication is entitled ‘Biochemical properties of ectoine hydroxylases from extremophiles and their wider taxonomic distribution among microorganisms’. It was published in ‘PLoS One’ (doi: 10.1371/journal.pone.0093809).

Widderich, N., Hoeppner, A., Pittelkow, M., Heider, J., Smits, S.H., and Bremer, E. Biochemical properties of ectoine hydroxylases from extremophiles and their wider taxonomic distribution among microorganisms. *PLoS One*. (2014). 9: e93809.

In this paper the phylogenetic distribution of ectoine biosynthetic genes was assessed and the salient biochemical and kinetic properties of six ectoine hydroxylases derived from different microorganisms were determined and compared to the already characterized EctD enzyme from *V. salexigens* (Bursy *et al.*, 2007). In addition, resolution of the crystal structure of EctD from the moderate halophile *Virgibacillus salexigens* in its apo-form is reported in this publication. This study provided the first detailed phylogenetic affiliation of ectoine biosynthesis, and contributed to the understanding of the biochemistry and crystal structure of the ectoine hydroxylase.

Personal contribution:
I conducted the data base searches and assessed the phylogenetic distribution of ectoine biosynthesis (together with Marco Pittelkow). I further performed overexpression and purification of the various proteins, determined their salient biochemical and kinetic properties (together with Marco Pittelkow) and analyzed the experimental data (together with the other authors). Figures and Tables were prepared by me (and Sander Smits). I also contributed to the writing of the manuscript together with Sander Smits as well as my PhD supervisor Erhard Bremer.

OPEN ACCESS Freely available online

PLOS ONE

Biochemical Properties of Ectoine Hydroxylases from Extremophiles and Their Wider Taxonomic Distribution among Microorganisms

Nils Widderich[1,2☯], Astrid Höppner[3☯], Marco Pittelkow[1], Johann Heider[1,5], Sander H. J. Smits[4]*, Erhard Bremer[1,5]*

1 Laboratory for Microbiology, Department of Biology, Philipps-University Marburg, Marburg, Germany, **2** Max Planck Institute for Terrestrial Microbiology, Emeritus Group R.K. Thauer, Marburg, Germany, **3** X-Ray Facility and Crystal Farm, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany, **4** Institute of Biochemistry, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany, **5** LOEWE-Center for Synthetic Microbiology, Philipps-University Marburg, Marburg, Germany

Abstract

Ectoine and hydroxyectoine are well-recognized members of the compatible solutes and are widely employed by microorganisms as osmostress protectants. The EctABC enzymes catalyze the synthesis of ectoine from the precursor L-aspartate-β-semialdehyde. A subgroup of the ectoine producers can convert ectoine into 5-hydroxyectoine through a region-selective and stereospecific hydroxylation reaction. This compatible solute possesses stress-protective and function-preserving properties different from those of ectoine. Hydroxylation of ectoine is carried out by the EctD protein, a member of the non-heme-containing iron (II) and 2-oxoglutarate-dependent dioxygenase superfamily. We used the signature enzymes for ectoine (EctC) and hydroxyectoine (EctD) synthesis in database searches to assess the taxonomic distribution of potential ectoine and hydroxyectoine producers. Among 6428 microbial genomes inspected, 440 species are predicted to produce ectoine and of these, 272 are predicted to synthesize hydroxyectoine as well. Ectoine and hydroxyectoine genes are found almost exclusively in *Bacteria*. The genome context of the *ect* genes was explored to identify proteins that are functionally associated with the synthesis of ectoines; the specialized aspartokinase Ask_Ect and the regulatory protein EctR. This comprehensive *in silico* analysis was coupled with the biochemical characterization of ectoine hydroxylases from microorganisms that can colonize habitats with extremes in salinity (*Halomonas elongata*), pH (*Alkalilimnicola ehrlichii*, *Acidiphilium cryptum*), or temperature (*Sphingopyxis alaskensis*, *Paenibacillus lautus*) or that produce hydroxyectoine very efficiently over ectoine (*Pseudomonas stutzeri*). These six ectoine hydroxylases all possess similar kinetic parameters for their substrates but exhibit different temperature stabilities and differ in their tolerance to salts. We also report the crystal structure of the *Virgibacillus salexigens* EctD protein in its apo-form, thereby revealing that the iron-free structure exists already in a pre-set configuration to incorporate the iron catalyst. Collectively, our work defines the taxonomic distribution and salient biochemical properties of the ectoine hydroxylase protein family and contributes to the understanding of its structure.

Citation: Widderich N, Höppner A, Pittelkow M, Heider J, Smits SHJ, et al. (2014) Biochemical Properties of Ectoine Hydroxylases from Extremophiles and Their Wider Taxonomic Distribution among Microorganisms. PLoS ONE 9(4): e93809. doi:10.1371/journal.pone.0093809

Editor: Beom Seok Kim, Korea University, Republic of Korea

Received December 14, 2013; **Accepted** March 6, 2014; **Published** April 8, 2014

Funding: Funding for this study was generously provided by grants from the Deutsche Forschungsgemeinschaft through the SFB 987 and the LOEWE program of the state of Hessen (via the Centre for Synthetic Microbiology, SYNMIKRO; Marburg) (both to J.H. and E.B.), a contribution from the Max-Planck Institute for terrestrial Microbiology (Marburg) through the Emeritus Group of R.K. Thauer (to N.W.), by the Fonds der Chemischen Industrie (to E.B.) and by the initiative "Fit for Excellence" of the Heinrich-Heine University of Düsseldorf (to A.H. and S.H.J.S.). N.W. war partly funded by the International Max Planck Research School for Environmental, Cellular and Molecular Microbiology (IMPRS-Mic, Marburg). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: sander.smits@hhu.de (SHGS); bremer@biologie.uni-marburg.de (EB)

☯ These authors contributed equally to this work.

Introduction

The ability to sensitively detect and respond in a timely manner to changes in the external osmolarity through concerted genetic and physiological adaptation reactions is critical for the wellbeing and growth of most microorganisms [1,2]. The accumulation of compatible solutes is a widely used strategy by members of both the *Bacteria* and the *Archaea* to offset the detrimental effects of high osmolarity on cellular hydration and physiology [3–5]. Compatible solutes are operationally defined as small organic osmolytes, highly water-soluble compounds whose physicochemical properties make them compliant with cellular biochemistry and physiology [6–9]. As a consequence, microbial cells can build-up compatible solute pools to exceedingly high intracellular levels, either through synthesis or uptake [1,4], and they do this in a manner that is sensitively tied to the degree of the environmentally imposed osmotic stress [10,11]. Accumulation of compatible solutes counteracts the efflux of water under hyperosmotic growth conditions; they thereby stabilize turgor and optimize the solvent properties of the cytoplasm [1,6,12]. These processes cooperate in strongly enhancing the growth of high osmolarity challenged cells.

Ectoine and its derivative 5-hydroxyectoine are well-recognized members of the compatible solutes [13,14] and are effective osmostress protectants for microorganisms [15,16]. Synthesis of ectoine proceeds from L-aspartate-β-semialdehyde and comprises three enzymatic steps that are catalyzed by L-2,4-diaminobutyrate transaminase (EctB), 2,4-diaminobutyrate acetyltransferase (EctA), and ectoine synthase (EctC) to yield the cyclic ectoine molecule [(4*S*)-2-methly-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid] [17,18]. The structural genes for the ectoine biosynthetic enzymes are typically organized in an operon (*ectABC*) [19] whose transcription is up-regulated in response to high osmolarity [11,20–25]. Enhanced transcription of the *ect* genes is also triggered in some microorganisms by extremes in growth temperature [21,26] as ectoines can also confer protection against both heat and cold stress [27–29]. A subgroup of the ectoine producers also synthesizes a hydroxylated derivative of ectoine, 5-hydroxyectoine [20,30], in a biosynthetic reaction that is catalyzed by the ectoine hydroxylase (EctD) [20,27,31].

In addition to their role in alleviating osmotic stress, ectoines also serve as stabilizers of macromolecules and even entire cells [15,32]. The function-preserving and anti-inflammatory effects of ectoines fostered substantial interest in exploring them for a variety of practical biotechnological applications and potential medical uses [15,32–34].

Despite their closely related chemical structures, 5-hydroxyectoine often possesses superior stress protecting and function preserving properties than its precursor molecule ectoine [29,35–38]. Here, we focus on the ectoine hydroxylase, the enzyme that forms (4*S*,5*S*)-2-methyl-5-hydroxy-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid from the precursor ectoine through a region-selective and stereospecific hydroxylation reaction [13,20]. The enzymatic characterization of the EctD protein from *Virgibacillus salexigens* [20] and *Streptomyces coelicolor* [29] identified the ectoine hydroxylase as a member of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily (EC1.14.11) [39–41]. The EctD-mediated hydroxylation of (4*S*)-ectoine to (4*S*,5*S*)-5-hydroxyectoine requires O_2 and 2-oxoglutarate as co-substrates, thereby forming CO_2, succinate, and 5-hydroxyectoine [20]. As seen in other members of the dioxygenase superfamily (e.g., the taurine dioxygenase TauD [42]), the EctD-catalyzed enzyme reaction is strictly dependent on a mononuclear ferrous iron center promoting the O_2-dependent oxidative decarboxylation of 2-oxoglutarate, a sequence of events coupled with a two-electron oxidation of the substrate ectoine [43]. The high-resolution (1.85 Å) crystal structure of the *V. salexigens* EctD enzyme [44] revealed a protein fold that is commonly observed in members of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily, the so-called jelly-roll or cupin fold [40,41]. The catalytically critical iron is coordinated by the side chains of a conserved H×D/E…H motive, the so-called 2-His-1-carboxylate facial triad [39–41].

To gain further insight into the properties of the ectoine hydroxylase and the taxonomic distribution of ectoine/hydroxyectoine producers, we have mined the genome sequences of members of the *Bacteria* and *Archaea* with fully sequenced genomes for the signature enzymes for ectoine (EctC) and hydroxyectoine (EctD) biosynthesis. We then explored the genome contexts of the *ect* gene clusters to identify those genes that are functionally associated with the production of ectoines, the specialized aspartokinase Ask_Ect [22,45] or with the genetic control of *ect* gene expression, the repressor protein EctR [24,25]. We coupled this comprehensive *in silico* analysis with the biochemical characterization of six EctD enzymes from phylogenetically widely separated bacteria covering various different lifestyles to define the properties and kinetic parameters of the ectoine hydroxylase on a broad basis. In addition, the crystal structure of the EctD protein from the salt tolerant moderate halophile *V. salexigens* in its iron-free form was solved, thereby allowing for the first time an assessment of the structural consequences of the binding of the active-site iron on the overall fold of the ectoine hydroxylase.

Results and Discussion

Database Searches for the Ectoine and Hydroxyectoine Biosynthetic Genes

To assess the prevalence and taxonomic distribution of the ectoine and hydroxyectoine biosynthetic genes in microorganisms, we searched through finished microbial genome sequences at the database of the U.S. Department of Energy (DOE) Joint Genome Institute [46] for the presence of an *ectC* ortholog, coding for the signature enzyme of the ectoine biosynthetic pathway, the ectoine synthase [19]. As a search query for this database analysis, we used the amino acid sequence of the *V. salexigens* EctC protein (accession number: AAY29688) [20]. At the time of the database search, 6428 microbial genomes were represented that were derived from 6179 members of the *Bacteria* and 249 members of the *Archaea*. Of these genomes, 440 contained an *ectC* gene (approximately 7%), and most of them were members of the *Bacteria*; the notable exceptions were five *ectC* sequences present in the genomes of *Archaea* (two *Methanosaeta* and three *Nitrosopumilus* species). Excluding closely related strains of the same species for our analysis and using only a single representative, we constructed a phylogentic tree of the EctC sequences (Fig. 1). It is apparent from our database analysis that ectoine is a compatible solute which is synthesized almost exclusively by members of the *Bacteria* (Fig. 1). Genome sequences of 139 strains of *Vibrio cholerae* are represented among the 6428 searched microbial genomes, each of which is predicted to produce ectoine, but only one of them was included in the dataset depicted in Fig. 1. The few predicted archaeal ectoine producers have probably acquired the ectoine biosynthetic genes via lateral gene transfer events, since the exchange of genetic material between members of the kingdoms of the *Bacteria* and *Archaea* is a well-documented phenomenon [47].

We then assessed the distribution of the ectoine hydroxylase orthologs (*ectD*) in bacterial and archaeal genomes by using the *V. salexigens* EctD protein (accession number: AAY29689) [20] as the search query to identify those microorganisms predicted to produce hydroxyectoine. We found that 272 of the sequenced genomes possessed an *ectD* gene. Invariably these microorganisms also possessed an *ectC* gene, a result that is expected from the fact that hydroxyectoine is synthesized directly from the precursor molecule ectoine [20]. Hence, about two-thirds of the putative ectoine producers are predicted to synthesize hydroxyectoine as well (Fig. 1). As expected from the oxygen-dependent reaction of the EctD enzyme, *ectD* is never present in genomes of obligate anaerobes, although it is not universally present in aerobic or facultative species. Consistently, from the above mentioned archaeal ectoine-producing representatives, only the three (aerobic) *Nitrosopumilus* species possess an *ectD* gene as part of their *ect* gene clusters, whereas the genome sequences of the two (anaerobic) *Methanosaeta* species lacked *ectD* altogether.

Overproduction and Purification of Recombinant Ectoine Hydroxylases from Extremophilic Bacteria

Biochemical properties of native ectoine hydroxylases from *V. salexigens* and *S. coelicolor* have been assessed previously [20,29]. To determine whether the reported features of these two studied EctD proteins are representative for ectoine hydroxylases in general, we

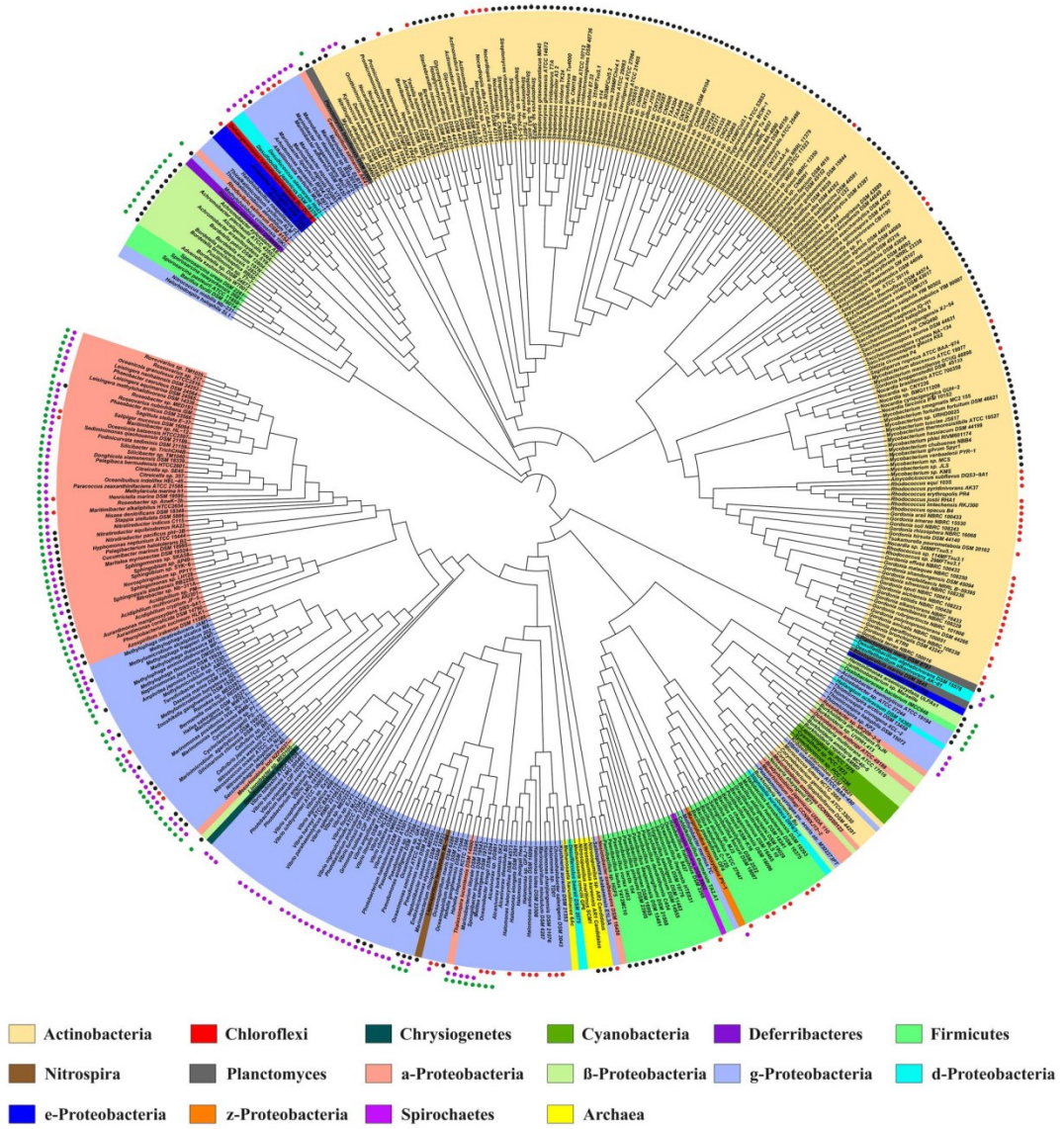


Figure 1. Phylogenetic tree of EctC- and EctD-type proteins. The shown phylogenetic tree is based on the alignment of EctC amino acid sequences identified by a BLAST search at the JGI Web-server that were then aligned using ClustalW. These compiled amino acid sequences were then used to assess the phylogenetic distribution of the EctC protein using the iTOL Web-server. Evolutionary distances are not given. The color code indicates the distribution of EctC among members of the *Bacteria* and *Archaea*. The presence of an *ectD* gene in a given microbial species possessing *ectC* is indicated by black (*ectD* is part of the *ect* gene cluster) or red circles (*ectD* is located outside of the *ect* gene cluster). Purple circles are indicating the presence of an *ask_ect* gene associated with the *ect* gene cluster, whereas the presence of an *ectR* regulatory gene is indicted by green circles. If different strains of the same species were sequenced, only one representative symbolizes them. For instance, there are genomic data of 139 strain of *Vibrio cholerae* available in the database, each of which possesses an *ectABC* gene cluster, but only one of these sequences was used for the phylogenetic analysis.
doi:10.1371/journal.pone.0093809.g001

set out to study the characteristics of this type of enzyme on a broader basis. For these biochemical studies we chose six EctD proteins from the following taxonomically widely separated and mostly extremophilic microorganisms: *Halomonas elongata*, *Acidiphi-*

lium cryptum, *Alkalilimnicola ehrlichii*, *Sphingopyxis alaskensis*, *Paenibacillus lautus*, and *Pseudomonas stutzeri*.

The Gammaproteobacterium *H. elongata* is the production strain for the industrial-scale manufacturing of ectoine [34] and grows in media with up to 5 M NaCl [48]. *A. cryptum* is an acidophilic metal-reducing Alphaproteobacterium that was isolated from an iron-rich sediment of an acid coal mine; it can grow at a pH of 5 [49,50]. *A. ehrlichii* is an arsenite-oxidizing haloalkaliphilic Gammaproteobacterium isolated from Mono Lake (CA, USA) and has a pH optimum of 9.3 [51]. The Alphaproteobacterium *S. alaskensis* is a cold-adapted marine ultra-microbacterium that was isolated from permanently cold (4–10°C) water sources in the Resurrection Bay (AK, USA) [52,53]. The Firmicute *P. lautus* was isolated from the Obsidian Hot spring in the Yellowstone National Park (WY, USA) that possesses a temperature range between 42–90°C; it can routinely be grown in the laboratory at 50°C [54]. The last studied microorganism was the nitrogen-fixing Gammabacterium *Pseudomonas stutzeri* strain A1501 that is not an extremophile, as it was isolated from plant roots [55]. Like the type strain of *P. stutzeri* (DSM 5109^T), it produces 5-hydroxyectoine very efficiently and in preference over ectoine [22,56], suggesting that its EctD enzyme might work particularly effectively.

Given the very different habitats of these microorganisms, we wondered if the biochemical properties of their EctD proteins would reflect the preferences of their producers with respect to the salt, pH, and temperature parameters prevalent in their natural habitats. Using the biochemically and structurally well characterized *V. salexigens* EctD protein (*Vs*EctD) [20,44] as a point of reference, the EctD proteins from the above-described six bacteria had an amino acid sequence identity ranging between 51% (*S. alaskensis*) and 40% (*H. elongata*). To study these EctD enzymes biochemically, we inserted the various *ectD* genes into an expression vector that allowed the production of the corresponding proteins as recombinant variants with a *Strep*-tag-II affinity peptide attached to their carboxy-terminus. These proteins could all be overproduced in an *Escherichia coli* host strain and isolated with good yields and purities by affinity chromatography on Strep-Tactin Superflow material (Fig. 2). The amino acid sequences of the native EctD proteins range in length between 302 and 306 amino acids, except for EctD of *H. elongata,* which is predicted to consist of 332 amino acids (Table 1). The migration of some of the purified recombinant EctD proteins on a 12% SDS-polyacrylamide gel (Fig. 2) deviates somewhat from their calculated molecular mass (Table 1), a property that might be connected with the particular amino acid composition of individual EctD proteins.

Since the presence of a correctly complexed iron ligand is critical for EctD-mediated enzyme catalysis [20,43,44], we determined the iron-content of each of these recombinant proteins and found between 0.87 and 0.96 mole iron per mol of EctD protein. Hence, these recombinant EctD proteins should all be functional. An initial assessment of their enzymatic activities under the same assay conditions as used previously for the ectoine hydroxylases from *V. salexigens* and *S. coelicolor* [20,29] demonstrated that this was indeed the case.

Biochemical Properties of the Ectoine Hydroxylases

We determined for each of the EctD enzymes its temperature and pH optimum and measured the influence of various salts (KCl, NaCl, K-glutamate, NH_4Cl) on the catalytic efficiency. The data from this set of experiments are summarized in Table 1 and are documented in detail for the *S. alaskensis* enzyme in Fig. 3. The data for all other enzymes are summarized in Fig. S1 to Fig. S5. Overall, the basic biochemical parameters of the six newly studied EctD enzymes and the re-analyzed EctD protein from *V. salexigens* [20] were all quite similar (Table 1), regardless of the environmental parameters that were prevalent in the habitats of those microorganisms from which they originate. However, differences were noted with respect to their resistance to the inhibiting action of increased salt concentrations (Table 1).

In studying the biochemical properties of the ectoine biosynthetic enzymes from *H. elongata*, Ono *et al.* [17] reported that the *in vitro* activity of these proteins was strongly dependent on high concentrations of NaCl (0.4–0.5 M), a type of salt that is unlikely to be accumulated to such high levels *in vivo* by osmotically stressed *H. elongata* cells, since sodium ions are toxic for bacterial cells. We did not find any strong stimulating effect of high NaCl concentrations on any of the ectoine hydroxylases we studied here (Table 1), including that of *H. elongata* (Fig. S1). On the contrary, high concentrations of NaCl typically inhibited the enzyme activities of the EctD variants (Fig. 3 and Fig. S1 to Fig. S5). However, notable stimulating effects [about two- to three-fold (Fig. 3 and Fig. S1 to Fig. S5)] on EctD enzyme activities were recorded with KCl or K-glutamate solutions.

We assessed the quaternary structure of the six newly studied EctD proteins by gel filtration. An example of this analysis is shown in Fig. S6 for the *S. alaskensis* EctD protein. The protein eluted between 72 to 83 ml (maximum: 77. 5 ml) from the size exclusion chromatography column and thereby corresponds to a protein of about 70.4 kDa. Since the calculated molecular mass of the *S. alaskensis* EctD protein monomer with the attached *Strep*-tag-II affinity peptide (nine amino acids) is 35.29 kDa, the ectoine hydroxylase is apparently a homodimer. The same conclusion was derived for all other analyzed EctD proteins (data not shown), including that from *V. salexigens*, which has previously been suggested to be a monomer [20].

Temperature Stability of the Ectoine Hydroxylases: The *S. Alaskensis* and *P. Lautus* Enzymes Stand Out

The studied EctD enzymes have similar temperature optima but differ in the range of temperatures in which they operate naturally (Table 1). To investigate this further, we studied their temperature stability. For these experiments, we pre-incubated 100 µg of each enzyme in 100 µl TES-buffer (pH 7–8) for 15 min at a given temperature and then measured its activity under assay and temperature conditions that had been optimized for each individual EctD protein (Table 1). The ectoine hydroxylase from *H. elongata* turned out to be the most temperature labile protein, whereas those from *S. alaskensis* and *P. lautus* proved to be quite temperature resistant; all other enzymes possessed intermediate degrees of temperature stability (Fig. 4). The strong temperature resistance of the *P. lautus* EctD protein does not come as a surprise since this *Paenibacillus* species was isolated from a hot spring with water temperatures ranging between 42–90°C [54]. The considerable heat tolerance of the *S. alaskensis* EctD enzyme is more of a surprise since this bacterium is well adapted to permanently cold (4–10°C) marine environments although it can grow at higher temperatures [52].

Kinetic Parameters of Ectoine Hydroxylases

After having optimized the parameters of the enzyme activity assays for each of the six purified ectoine hydroxylases (Table 1), we determined their apparent kinetic parameters for the co-substrate 2-oxoglutarate and the substrate ectoine (Table 2). This assessment showed that the studied ectoine hydroxylases all possess similar kinetic parameters. For instance, the *S. alaskensis* enzyme had an apparent K_m of 9.8±0.5 mM for its substrate ectoine and of

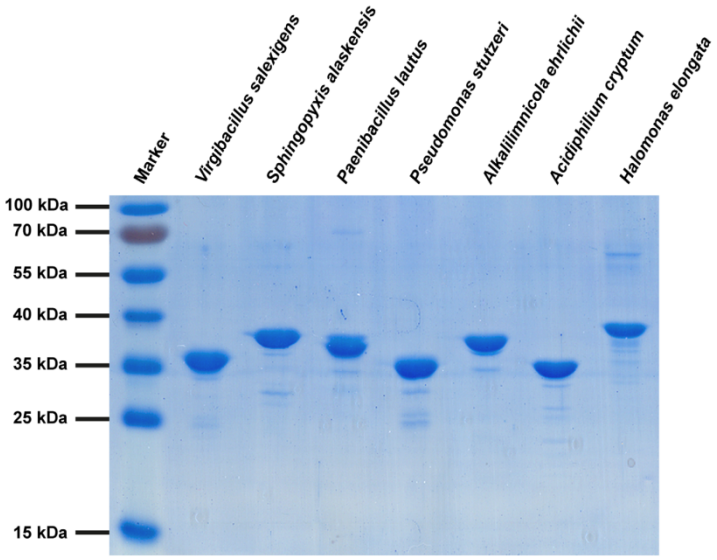


Figure 2. Purification of recombinant EctD proteins. A 12% SDS-PAGE of the recombinant EctD-type proteins originating from different microbial species after their overproduction in *E. coli* and purification via *Strep*-tag-II affinity chromatography is shown. 5 µg of each purified EctD protein were applied onto the gel. The gel was run at 25 mA for 2.5 h. The PageRuler Prestained Protein Ladder (Thermo Scientific, Schwerte, Germany) was used as marker. We note the presence of a overlapping second band in protein sample of the *Paenibacillus lautus* EctD preparation. This might stem from overloading the gel somewhat or from partial degradation of the purified EctD protein.
doi:10.1371/journal.pone.0093809.g002

2.7±0.3 mM for its co-substrate 2-oxoglutarate, respectively, a V_{max} of 1.0±0.2 U mg^{-1}, a k_{cat} of 1.2 s^{-1} per holoenzyme and a catalytic efficiency of 0.12 mM^{-1} s^{-1} (Table 2). The *P. stutzeri* and *V. salexigens* enzymes stand out among the tested enzymes with respect to their catalytic efficiencies with values of 1.44 mM^{-1} s^{-1} and 1.31 mM^{-1} s^{-1}, respectively (Table 2). In contrast, the EctD enzyme from the industrially used ectoine/hydroxyectoine production strain *H. elongata* [32,34] exhibits no particularly notable features with respect to its catalytic efficiency (0.49 mM^{-1} s^{-1}) (Table 2). The relatively good performance of the ectoine hydroxylase from *P. stutzeri* A1501 is certainly consistent with the preferred accumulation of 5-hydroxyectoine by osmotically stressed cells of this isolate over that of ectoine [22], a feature that is also found in the type strain (DSM 5190^T) of *P. stutzeri* [56].

Our data show that ectoine hydroxylases are not particularly effective enzymes since their affinities for their substrate ectoine and their co-substrate 2-oxoglutarate are low with apparent K_m values in the mM range, and they exhibit only modest V_{max} numbers and catalytic efficiencies (Table 2). These properties of EctD enzymes have also been observed previously when the ectoine hydroxylases from *V. salexigens*, and *S. coelicolor* were isolated as native proteins from their natural producer bacteria and not as recombinant proteins as done here [20,29]. The moderate kinetic parameters of the ectoine hydroxylase might be connected to the fact that in osmotically stressed microbial cells, ectoine is typically accumulated first and 5-hydroxyectoine production then sets in only after a substantial cellular pool of its precursor molecule has been built up [22,29,57]. Furthermore, the *in vitro* activities of the EctD enzymes require considerable 2-oxoglutarate concentrations in order to work efficiently (Table 2) [20,29], and therefore the cellular 2-oxoglutarate pool [58] could potentially limit 5-hydroxyectoine formation *in vivo*.

EctD Enzyme Activity is not Reversible

In their recent excellent overview on the role of ectoines as microbial stress protectants and their biotechnological applications, Pastor *et al.* [15] suggested that the EctD enzyme may also catalyze the reverse reaction to form ectoine from 5-hydroxyectoine (see Fig. 2 in [15]), albeit without providing any experimental evidence or presenting a possible mechanism. We therefore assayed for the stability of the products of the enzyme reaction of the *V. salexigens* EctD protein under conditions set up to favor a hypothetical backward reaction (6 mM 5-hydroxyectoine, 10 mM bicarbonate, and 20 mM succinate as potential substrates were incubated with 40 µg EctD protein) and found no decrease of the hydroxyectoine concentration or production of any ectoine, even after incubating the enzyme reaction mixture for 24 hours (Fig. S7). For comparison, 6 mM ectoine are almost completely converted into 5-hydroxyectoine within 20 min when 10 mM 2-oxoglutarate were provided as the co-substrate (Fig. S7). This is also predicted from the highly exergonic thermodynamics of ectoine hydroxylation by EctD (estimated $\Delta G^{\circ\prime} < -400$ kJ/mol), which should completely preclude the backward enzyme reaction [39–41,59]. We therefore conclude that the ectoine hydroxylase is an enzyme that operates exclusively in one direction under physiologically relevant conditions to direct the formation of 5-hydroxyectoine from the precursor ectoine.

Crystal Structure of the *V. Salexigens* EctD Protein in its Iron-free Form

A high-resolution (1.85Å) crystal structure of the EctD protein from *V. salexigens* has previously been reported in complex with the catalytically important iron ligand; however, it lacks the co-substrate 2-oxoglutarate and the substrate ectoine [44]. This structure was recently used as a starting point for molecular

Table 1. Biochemical properties of the studied EctD-type proteins.

EctD from	length [AS]	mass [kDa]	pI	optimum temp. [°C]	temp. range [°C]	pH	pH range	optimum KCl [mM]	KCl range [mM]	optimum NaCl [mM]	NaCl range [mM]
V. salexigens	300	34.4	5.8	32	5–50	7.5	5.5–9.6	150	0–750	100	0–350
S. olaskensis	306	34.1	5.5	40	5–50	8.0	5.5–9.6	100	0–1000	100	0–500
H. elongata	332	37.4	5.8	32	5–47	8.0	6.5–9.6	150	0–750	100	0–250
P. stutzeri	302	34.2	5.5	35	10–50	7.5	5.5–9.6	150	0–1000	100	0–350
P. lautus	302	34.8	5.6	40	15–50	7.5	5.5–9.6	200	0–750	150	0–250
A. ehrlichii	302	34.3	5.7	35	15–45	7.5	6.5–9.6	150	0–1000	150	0–400
A. cryptum	306	34.1	5.8	32	10–47	8.0	5.5–9.6	100	0–1000	50	0–300

The biochemical properties of the studied EctD-type proteins were determined as described in Material and Methods. The given temperature, pH and salt ranges describe a window in which the tested enzymes still exhibited some degree of activity.

doi:10.1371/journal.pone.0093809.t001

dynamics simulations and site-directed mutagenesis experiments to glean information about the coordination of the ligands within the EctD active site [43]. We continued our efforts to obtain an EctD crystal structure containing all ligands and therefore pushed the recombinant production of the *V. salexigens* EctD protein in *E. coli* to very high levels in order to supply the large quantities of protein needed for the crystallization trials. In this way, we increased the amounts of the purified *V. salexigens* recombinant EctD enzyme from about 20–25 mg per liter of culture (the ectoine hydroxylase source for biochemical studies) to 200–300 mg per liter of culture. However, after analyzing the iron content of this strongly overproduced EctD enzyme preparation, it became apparent that most of the isolated proteins did not contain an iron molecule; the iron content of the EctD protein solution dropped to 0.1–0.2 mole iron per 1 mol of EctD, rendering the enzyme largely inactive. Upon addition of Fe^{2+} ions prior to the enzyme activity measurements, we observed that the activity returned to levels observed before [20,43], indicating that the missing iron catalyst in the purified EctD protein can be restored after the protein has adopted its native cupin barrel fold [44].

These observations prompted us to explore whether the *V. salexigens* EctD protein adopts a similar conformation in its iron-free and iron-bound forms, or whether the incorporation of the iron ligand leads to substantial structural changes. We grew EctD crystals and collected a 1.9 Å X-ray dataset. The cell constants and the space-group (Table S1) were identical to the structure of the iron-bound EctD, suggesting that the apo-EctD protein crystalized in a manner similar to that found in the iron bound form [44]. After solving the new crystal structure of the EctD protein, it became apparent that the iron ligand was lacking, as evidenced by the missing pronounced electron density that is present in the iron-bound EctD crystal structure [44]. Otherwise, the apo- and the iron-bound forms are almost identical, as indicated by the RMSD value of 0.34 Å over 280 Cα atoms. An overlay of both EctD crystal structures is shown in Fig. 5A.

In the *V. salexigens* EctD protein, the iron ligand is bound via interaction with two histidine side-chains, His-146 and His-248, and the side-chain of Asp-148 (Fig. 5A and B) [44]. Together these residues form a conserved H×D/E…H motif, the so-called 2-His-1-carboxylate facial triad. [39–41,59]. A comparison of the iron-binding residues in the apo- and iron-bound structures of the *Vs*EctD protein shows that they exhibit the same architecture, except that the iron ligand is present in one structure and absent in the other (Fig. 5B). Interestingly, in the apo-structure of EctD, two water molecules populate the iron-binding site formed by the 2-His-1-carboxylate facial triad. This keeps the side chains of the His-146, His-248 and Asp-148 in an orientation very similar to that observed in the iron-bound EctD crystal structure (Fig. 5B). Hence, the EctD apo-protein exists in a form that is pre-set to incorporate the iron catalyst [43].

Phylogenetic Distribution of the *EctC* and *EctD* Genes

Previous studies have indicated that the ability to produce ectoine and hydroxyectoine is widely distributed in the microbial world but is absent from eukarya [15,34,44,60]. We updated and extended this information on a genome-wide scale in the following way: (i) first, we visualized the relationship among the 440 retrieved EctC sequences via the iTOL tool [61] to analyze their taxonomic association with members of the *Bacteria* and *Archaea*; (ii) we then projected the information on the presence of the ectoine hydroxylase in a given microbial species onto this phylogenetic tree of the EctC protein to reveal the extent and taxonomic distribution of EctD protein among putative ectoine producers; (iii) we inspected the genome context of each of the 440 microbial

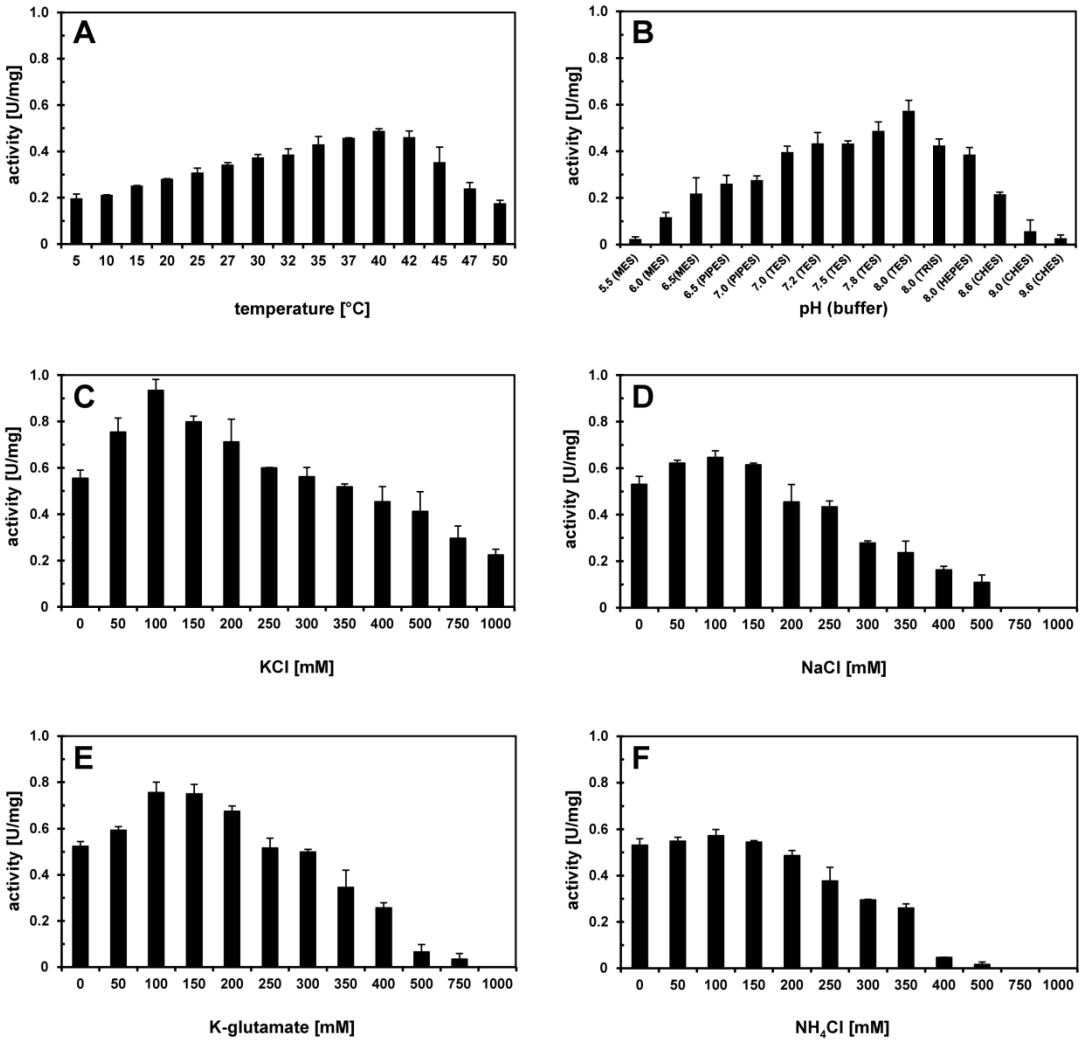


Figure 3. Biochemical properties of the EctD enzyme from *S. alaskensis*. The enzyme activity of the ectoine hydroxylase from *S. alaskensis* is shown with respect to (A) the temperature optimum, (B) the pH optimum and the influence of different salts: (C) potassium chloride, (D) sodium chloride, (E) potassium glutamate and (F) ammonium chloride.
doi:10.1371/journal.pone.0093809.g003

species that possessed *ectC* and from this bioinformatics approach retrieved the genetic organization of the *ect* biosynthetic gene cluster; (iv) furthermore, we assessed the co-localization of the *ect* genes with genes that have been functionally associated with ectoine/hydroxyectoine biosynthesis, the gene for a specialized aspartokinase Ask_Ect [22,45,60], and that of the transcriptional regulator EctR [24,25].

In the first step of this *in silico* analysis, we aligned the retrieved 440 EctC sequences using the ClustalW [62] algorithm and found amino acid sequence identities that ranged between 88% and 27% with reference to the *V. salexigens* EctC protein [20]. The corresponding numbers for the degree of identity of the 272 EctD proteins range between 79% and 37% with reference to the *V. salexigens* EctD protein. The visualization of the taxonomic distribution of the EctC and EctD proteins with the iTOL-software package [61] revealed putative bacterial and archaeal ectoine producers in 17 phyla (Fig. 1). Fifteen of these phyla are taxonomically associated with the domain of the *Bacteria* and two with the domain of the *Archaea*. The taxonomic distribution of the putative hydroxyectoine producers was more restricted: ectoine hydroxylase genes are found only in nine phyla (Fig. S8). In the following, we further consider the ectoine synthase, the ectoine hydroxylase, the specialized aspartokinase Ask_Ect, and the transcriptional regulator EctR.

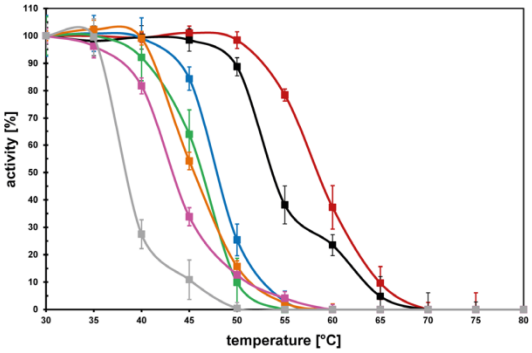


Figure 4. Resistance of various ectoine hydroxylases against the denaturing effects of high temperature. The temperature profiles of the ectoine hydroxylases from *H. elongata* (grey), *A. cryptum* (pink), *A. ehrlichii* (orange), *V. salexigens* (green), *P. stutzeri* (blue), *S. alaskensis* (black), and *P. lautus* (red) are given. Each EctD protein was pre-incubated at the indicated temperatures for 15 min before its specific activity was then determined under its optimal assay condition. The enzyme activity exhibited by each enzyme after pre-incubation at 30°C was set as 100%.
doi:10.1371/journal.pone.0093809.g004

The Ectoine Synthase EctC

The ectoine biosynthetic enzymes L-2,4-diaminobutyrate transaminase (EctB; EC 2.6.1.76) and the 2,4-diaminobutyrate acetyltransferase (EctA; EC 2.3.1.178) have isoenzyme counterparts in various biochemical pathways [17,18], but the ectoine synthase (EctC; EC 4.2.1.108) is unique. Therefore, EctC has been considered so far as the diagnostic enzyme for ectoine production. It catalyzes the cyclization of *N*-γ-acetyl-2,4-diaminobutyrate to ectoine via a water elimination reaction [17,18], whereas *N*-α-acetyl-2,4-diaminobutyrate that is formed during ectoine catabolism [34] is apparently no substrate for this enzyme. In a reversal of the native cyclization reaction scheme, EctC can also inefficiently hydrolyze synthetic ectoine derivatives with reduced or expanded ring sizes and can catalyze the cyclic condensation of glutamine to the synthetic compatible solute 5-amino-3,4-dihydro-2H-pyrrole-2-carboxylate (ADPC) as a side reaction [57].

The distribution of the ability to synthesize ectoine, as indicated by the presence of an *ectC* gene in a given genome sequence, appears to extend mostly to members of the Proteobacteria, Firmicutes, or Actinobacteria (Fig. 1). Moreover, the EctC sequences cluster in large part with the taxonomic subgroups of these bacterial phyla and their branching order occurs in parallel with the different corresponding taxonomic units of their microbial hosts down to the order level. This suggests a long co-evolution of the ectoine synthesis genes in the various bacteria. Moreover, the analysis revealed a conspicuous abundance of marine species among the putative proteobacterial ectoine producers. On the other hand, the ectoine-producing species of the Firmicutes and Actinobacteria are mostly from terrestrial habitats. The few EctC sequences from species with different taxonomic affiliations are interspersed into large clusters of related species (Fig. 1) and can probably be explained by lateral gene transfer events. Thirteen amino acid residues were fully conserved among all of the 440 inspected EctC proteins, but in contrast to EctD [20,44], no signature sequence of the ectoine synthase was readily discernable.

In our dataset of 440 putative ectoine producers, we found 22 organisms containing solitary genes for EctC-type proteins that were not associated with the characteristic *ectA* and *ectB* ectoine biosynthetic genes [19] (Fig. 6). They all possess genes for EctB-related proteins somewhere else in their genome, but none of them possess genes for recognizable EctA-like proteins. Notably, the *ectD* gene is also completely absent from this group of bacteria (Fig. 1). Kurz and co-workers [63] investigated the ectoine biosynthetic potential of the plant pathogen *Pseudomonas syringae* pv. *syringae* B728a, a strain with such an orphan *ectC* gene, and found that it produced ectoine only under osmotic stress conditions when surface-sterilized leaves of the host plant *Syringa vulgaris* were added to the bacterial culture. Furthermore, in functional complementation experiments, the corresponding ectoine synthase was only partially active and, surprisingly, feeding of the direct ectoine precursor, *N*-γ-acetyl-2,4-diaminobutyrate, to the *P. syringae* pv. *syingae* B728a did not lead to ectoine synthesis although this compound was taken up by the cells [63].

Inspection of the EctC phylogenetic tree (Fig. 1) showed that the 22 host species possessing these solitary *ectC* genes are taxonomically rather diverse. The retrieved amino acid sequences are all phylogenetically related to a cluster of EctC proteins present primarily in members of the Firmicutes that all possess intact ectoine biosynthetic pathways and that are predicted to produce both ectoine and hydroxyectoine (Fig. 1). In our view, the functional relationships of these solitary *ectC* genes cannot yet be fully determined with confidence: (i) the species possessing orphan EctC-type proteins may be actual ectoine producers that have to rely on an environmental supply of ectoine precursor molecules as suggested by the data reported by Kurz *et al.* [63]; (ii) these EctC-like proteins may be evolutionary remnants of a previously intact

Table 2. Kinetic parameters of the analyzed ectoine hydroxylases.

EctD from	K_m [mM ectoine]	v_{max} [U/mg]	k_{cat} [s^{-1}]	k_{cat}/K_m [$mM^{-1}\ s^{-1}$]	K_m [mM 2-oxoglutarate]
V. salexigens	5.9±0.3	6.4±0.2	7.7	1.31	4.9±0.3
S. alaskensis	9.8±0.5	1.0±0.2	1.2	0.12	2.7±0.3
H. elongata	5.7±0.6	2.5±0.2	2.8	0.49	4.8±0.4
P. stutzeri	6.2±0.4	6.7±0.2	8.9	1.44	4.6±0.5
P. lautus	9.5±0.7	1.3±0.1	1.6	0.17	3.9±0.2
A. ehrlichii	9.0±0.3	1.0±0.1	1.2	0.13	5.0±0.3
A. cryptum	10.0±0.6	2.8±0.3	3.4	0.34	4.1±0.4

The kinetic parameters of the studied EctD enzymes were determined under conditions that were optimal for each enzyme (see Table 1) by independently varying the substrate concentration of ectoine between 0 and 40 mM and that of the co-substrate 2-oxoglutarate between 0 and 50 mM. The k_{cat} values were determined per holoenzyme (a homo-dimer of the EctD protein) and the catalytic efficiency for the hydroxylation of ectoine is given as k_{cat}/K_m.
doi:10.1371/journal.pone.0093809.t002

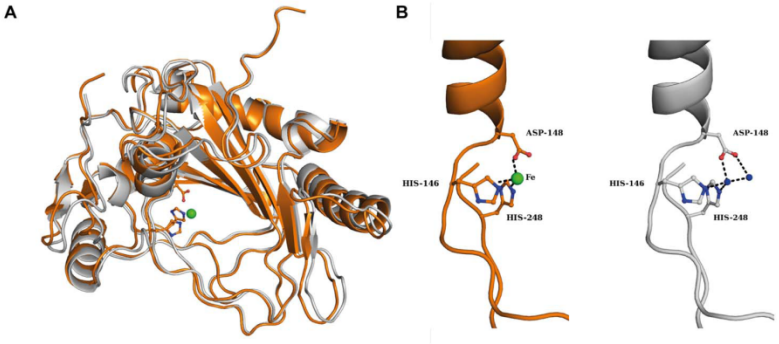


Figure 5. Crystal structure of the apo-form of the ectoine hydroxylase from *V. salexigens*. (A) Overlay of the crystal structure of the apo-EctD protein (colored in grey) with the Fe-bound crystal structure of EctD (colored in orange) in cartoon representation. The Fe ion of the Fe-bound EctD protein is represented as a green sphere. Data coordinates for the iron-bound form of the *V. salexigens* EctD protein were taken from the protein database (PDB) entry 3EMR and those from the iron-free form were from PDB entry 4NMI. (B) Details of the molecular determinants of the iron-binding site of the *V. salexigens* EctD protein in its iron bound (orange) and iron-free (grey) forms. The side chains of the iron-binding residues Asp148, His146 and His248 are highlighted. Green and blue spheres represent the bound iron and water molecules, respectively.
doi:10.1371/journal.pone.0093809.g005

ectoine biosynthetic pathway; or (iii) may have evolved (or be in the process of evolving) towards biochemical activities other than the cyclization of the direct ectoine precursor molecule *N*-γ-acetyl-2,4-diaminobutyrate.

The Ectoine Hydroxylase EctD

The ectoine hydroxylase [20,27,31] is frequently confused in genome annotations with proline- or phytanoyl-hydroxylases that, like EctD, also belong to the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily (EC1.14.11) [39–41]. However, *bona-fide* EctD-type proteins can be distinguished from the latter two enzymes by the presence of a strictly conserved signature sequence [20,44]. This stretch of 17-amino acids [Phe-143 to Pro-159 in the *Vs*EctD protein] not only serves an important role for the structuring of the overall fold of the EctD cupin barrel [44], but it also contains a number of residues implicated by *in silico* modeling and by structural and mutational analysis in the binding of iron, 2-oxoglutarate and ectoine [43].

ectD genes are only present in a subset of strains that also possess the full set of ectoine biosynthetic genes but the extent of *ectD* occurrence varies widely between different bacterial taxa. The *ectD* gene can either be part of the *ectABC* operon or can be encoded somewhere else in the genome [20,27,31] (Fig. 6). Of the putative 272 hydroxyectoine producers, 72% possess *ectD* genes that are located next to the *ectABC* gene cluster (Fig. 1 and Fig. 6). As expected from the oxygen dependence of the ectoine hydroxylase enzyme reaction [20,43], none of the obligately anaerobic ectoine-producing bacterial or archaeal species contains an *ectD* gene. The highest incidence of *ectD* is observed in the actinobacterial ectoine producers, in particular in all sequenced species of the genera *Streptomyces* and *Actinomadura*, all fast-growing (but none on the slow-growing) species of *Mycobacterium*, and most sequenced species of the orders *Pseudonocardiales*, *Glycomycetales*, the genus *Nocardiopsis* and the phylogenetically basal genus *Nitriliruptor* (Fig. 1 and Fig. S8). It is noteworthy in this context that the ectoine hydroxylase from *S. coelicolor* has been biochemically characterized [29] and it exhibits kinetic parameters similar to the six ectoine hydroxylases that were functionally assessed in this study (Table 2). There are only three species belonging to the *Archaea* that are predicted to synthesize hydroxyectoine. They are all members of the aerobic thaumarchaeal genus *Nitrosopumilus* [64], and they share very similar *ectD* gene products with those of the gammaproteobacterial genus *Nitrosococcus*. As both genera represent marine nitrifying microorganisms, recent gene sharing by lateral gene transfer [47] seems quite plausible.

The Specialized Aspartokinase Ask_Ect

Ectoine is formed from L-aspartate-β-semialdehyde [17,18], a central hub in microbial amino acid metabolism, cell wall biosynthesis, and antibiotic production [65]. L-aspartate-β-semialdehyde is synthesized from L-aspartate through the subsequent enzymatic activities of the aspartokinase (Ask) and the aspartate-semialdehyde-dehydrogenase (Asd). Complex transcriptional and posttranscriptional control mechanisms directed towards the expression of the *ask* gene and the enzymatic activity of its encoded protein ensure that there is no over- or undersupply of L-aspartate-β-semialdehyde [65].

The cellular pool of L-aspartate-β-semialdehyde is a potential bottleneck for the massive ectoine synthesis setting in under high osmolarity growth conditions [66]. To avoid such a metabolic constraint, a sub-group of the ectoine producers increase the cellular level of a specialized aspartokinase, Ask_Ect, simultaneously with the amounts of the EctABC biosynthetic enzymes through the co-transcription of the corresponding structural gene (*ask_ect*) with the *ectABC/D* gene cluster [22,45]. However, it is apparent from our database analysis that the majority of the ectoine/hydroxyectoine producers possibly circumvent such an anabolic bottleneck without producing an aspartokinase that is specifically earmarked for ectoine production. We found that about 30% (132 microbial species) of the predicted 440 ectoine/hydroxyectoine producers possess an *ask_ect* gene. This gene is almost exclusively found in Alpha-, Gamma- and Deltaproteobacteria; the taxonomic distribution among the Gammaproteobacteria is somewhat patchy but it is quite regular in the other two subphyla and the gene seems to be absent in Betaproteobacteria (Fig. 1). Strikingly, none of the predicted ectoine/hydroxyectoine producers that belong either to the Firmicutes or to the Actinobacteria (except three species) possess an *ask_ect* gene in the vicinity of their *ectABC/D* gene clusters (Fig. 1).

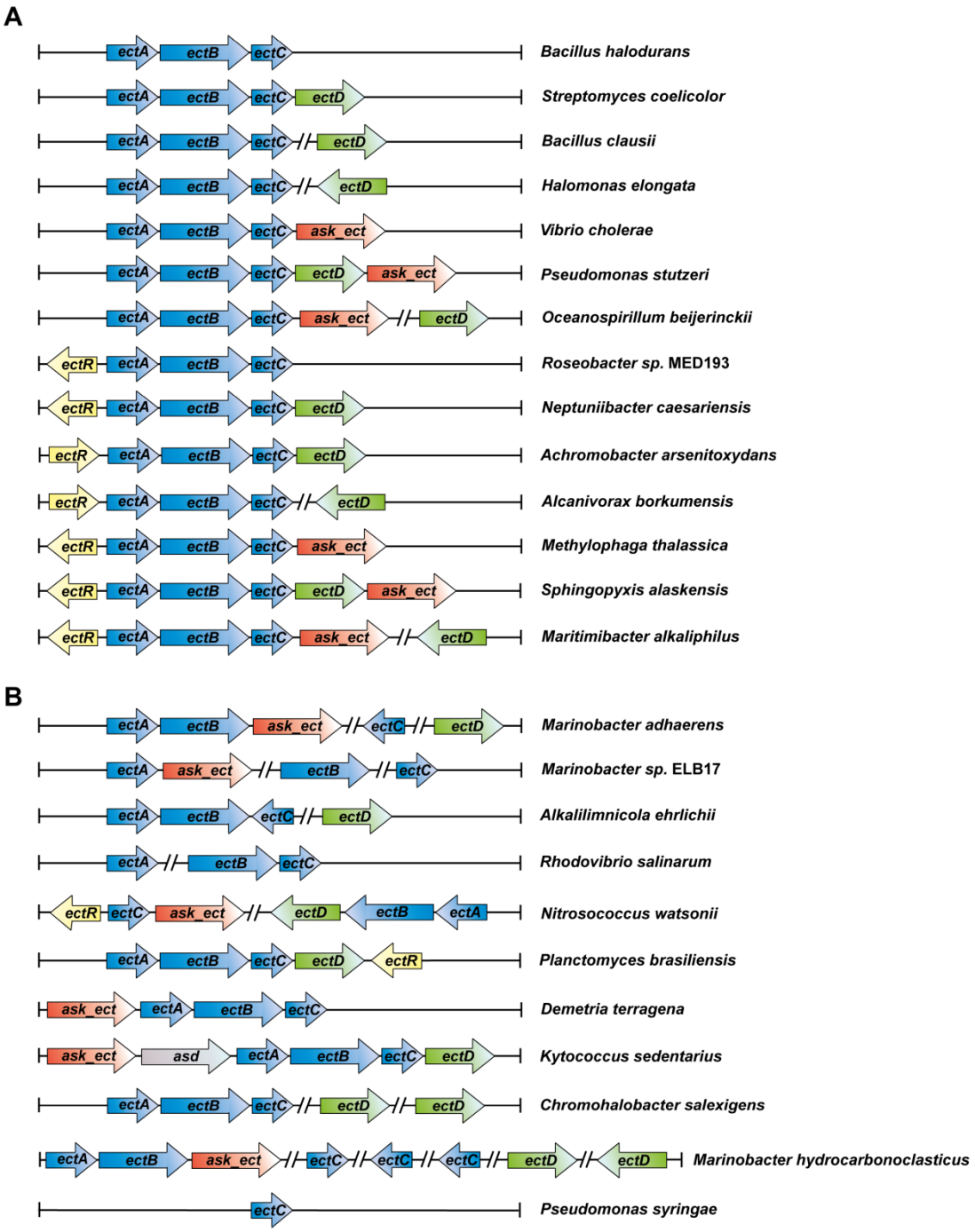


Figure 6. Genetic organization of the ectoine/hydroxyectoine biosynthesis gene clusters. The different types of *ect* gene clusters present in putative ectoine/hydroxyectoine producers are represented. An example for the genetic organization of each type of *ect* cluster found in the *ectC* reference set (440 representatives; Fig. 1) is given along with a microorganism in which it occurs. (A) Most common organizational types of the *ect*

gene clusters. (B) Representatives of the organization of the ectoine/hydroxyectoine biosynthetic genes that deviate from the otherwise commonly found genetic organization.
doi:10.1371/journal.pone.0093809.g006

Aspartokinases are ubiquitously found in microorganisms and several aspartokinases with distinct regulatory features are often present in the same bacterial cell [65]. The latter is true for the ectoine/hydroxyectoine producer *Pseudomonas stutzeri* A1501, where a comparative biochemical analysis of the specialized aspartokinase Ask_Ect and the anabolic aspartokinase LysC revealed distinct feedback inhibition profiles by metabolites [22].

The Transcriptional Regulator EctR

The functional association of the *ectR* gene with ectoine biosynthesis was first demonstrated in the halotolerant methanotrophic Gammaproteobacterium *Methylomicrobium alcaliphilum*, where EctR serves as a repressor of *ectABC-ask_ect* gene transcription [25]. Notably, the elevated transcription of the ectoine biosynthetic genes in an *ectR* mutant of *M. alcaliphilum* remains osmotically inducible [25]. It is currently not known which environmental or cellular cues dictate the binding to or the release of EctR from its operator sequence. In *M. alcaliphilum*, the EctR operator overlaps the −10 sequence of the *ect* promoter and EctR might also regulate the expression of its own structural gene [25]; however, this latter regulatory feature does not always seem to exist [24]. The EctR repressor protein is a member of the widely distributed group of MarR transcriptional regulators but forms a distinct sub-group within this superfamily [24]. Of the 440 putative ectoine/hydroxyectoine producers, 24% possess *ectR*-type genes (107 microbial species) (Fig. 1) whose transcriptional direction is frequently oriented divergently from that of the *ect* gene cluster (Fig. 6). EctR is found almost exclusively among the Proteobacteria; all Alpha- and Betaproteobacteria that are predicted to synthesize ectoines possess an *ectR* gene, whereas its distribution among the Gammaproteobacteria is more irregular (Fig. 1).

It is worth noting that in *Vibrio cholerae* an EctR-related MarR-type transcriptional regulator (CosR) has been described that negatively controls ectoine biosynthetic and compatible solute uptake genes in response to the ionic strength of the growth medium. CosR from *V. cholerae* and EctR from *M. alcaliphilum* exhibit 51% amino acid sequence identity; however, unlike *ectR*, the *cosR* gene is not located in the vicinity of the ectoine gene cluster present in *V. cholerae* [67]. Ectoine biosynthesis, but not that of hydroxyectoine, is widespread among *V. cholerae* strains [68] and other *Vibrio* species [69], microorganisms that primarily live in marine habitats and estuarine ecosystems. For instance, genome sequence of 139 *V. cholerae* strains have been deposited in the databases and each of these strains is predicted to synthesize ectoine (data not shown).

Genetic Organization of the Ectoine and Hydroxyectoine Structural Genes

After the initial discovery of the *ectABC* gene cluster for the synthesis of ectoine in *Marinococcus halophilus* [19], transcriptional profiling of the corresponding genes in several Gram-negative and Gram-positive bacteria showed that they were transcribed as operons inducible by osmotic or temperature stress [11,20,21,23,26,27]. In some ectoine producers, the *ectABC* genes are expressed from a single promoter [11,20], whereas in others a more complex pattern of regulation of this gene cluster has been reported [21,24,34]. Evidence for the presence of a putatively nitrogen-responsive Sig-54 type promoter driving the separate expression of *ectC* has been presented in the case of *H. elongata* [34].

We inspected the genetic organization of the *ect* genes in the 440 microbial species that we regarded as putative ectoine producers from our database analysis (Fig. 1). In the vast majority (85%), the *ectABC* genes were located next to each other, strongly suggesting that their transcriptional organization is centrally based on an evolutionarily highly conserved operon structure (Fig. 6A). The basic *ectABC* gene cluster is frequently associated either with *ectD*, *ask_ect*, or *ectR*, and various genetic configurations of these *ect*-associated genes can be found in microbial genomes (Fig. 6A). The *ectD* gene may either be part of the *ectABC* operon or form a separate transcription unit somewhere else in the genome.

Our database analysis shows that the genetic organization of the *ect* gene cluster is well preserved in groups of microorganisms that are widely separated taxonomically (Fig. 6A). Nevertheless, there is a substantial sub-group (about 15%) of the putative ectoine/hydroxyectoine producers where the ectoine/hydroxyectoine biosynthetic genes and the functionally associated *ask_ect* and *ectR* genes are not organized in the well-defined gene clusters found in 85% of our reference data set. In this group of bacteria, the order of the various *ect* genes have either been scrambled or they have been separated from each other on the chromosome (Fig. 6B). It is currently unclear whether this non-canonical gene organization has any consequences for transcriptional induction or the level of ectoine/hydroxyectoine production in response to osmotic or temperature stress since, to the best of our knowledge, none of the putative ectoine/hydroxyectoine producers with deviating gene organizations have been functionally studied.

In a few of the hydroxyectoine producers (e.g., *Arthrobacter castelli* DSM 16402, *Marinobacter aquaeolei* VT8, *Rhodococcus opacus* B4, *Rhodococcus* sp. RHA1, *Chromohalobacter salexigens* DSM 3043), two copies of the *ectD* gene are found (Fig. 6B). Studies with *C. salexigens* DSM 3043 have shown that only one of these ectoine hydroxylases is responsible for the production of the majority of the hydroxyectoine found in this highly salt-tolerant bacterium [70]. In the genomes of other microorganisms, several *ectC*-type genes are found (e.g.; there are three *ectC* genes in *Marinobacter aquaeolei*). However, nothing is known whether these genes are all expressed and what (if any) the functional consequences of multiple, closely related EctC proteins within the same bacterium might be.

Another interesting finding of our database analysis is the identification of microorganisms that possess either two (e.g., *Phaeobacter arcticus* DSM 23566, *Vibrio cholerae* 0395, *Streptomyces flavogriseus* ATCC 33331) or even three (e.g., *Streptomyces clavuligerus* ATCC 27064) full copies of the ectoine/hydroxyectoine biosynthetic gene cluster. A pairwise comparison of the amino acid sequence of the various Ect proteins within a given species indicates that the increase in the *ect* gene copy number has likely arisen via gene duplication events since the corresponding proteins are all closely related to each other (data not shown). Whether these extra copies of the ectoine/hydroxyectoine biosynthetic genes are actually functionally expressed and whether the bacteria with the additional *ect* gene copies produce more ectoine than those with only one copy of the corresponding genes remains an interesting question for future studies.

Concluding Remarks

Our comprehensive database analysis of finished microbial genomes revealed at the time of the search (July 2013) the presence

of ectoine biosynthetic genes [19] in about 7% of the represented microorganisms, and about two thirds of these are predicted to produce hydroxyectoine [20,27,31] as well. Since hydroxyectoine often possess function-preserving and stress-relieving properties superior to those of ectoine [29,35–38], one wonders why not all ectoine producers synthesize 5-hydroxyectoine since this can readily be accomplished from ectoine in a single step [20]. Part of the answer to this question becomes apparent when one considers the physiological requirements and the oxygen dependence of the ectoine/hydroxyectoine producer microorganisms. The EctD-catalyzed formation of 5-hydroxyectoine is an O_2-dependent enzyme reaction [20,43] and consequently none of the predicted hydroxyectoine producers is an obligate anaerobe, whereas both aerobic and anaerobic microorganisms can produce ectoine (Fig. 1).

Our data view the potential ectoine and hydroxyectoine producers within a wider taxonomic context (Fig. 1 and Fig. S8). With a few notable exceptions that revealed ectoine/hydroxyectoine biosynthetic genes in five members of the *Archaea* (*Methanosaeta* and *Nitrosopumilus* species), ectoine and hydroxyectoine producers are taxonomically affiliated with the domain of the *Bacteria*. We assessed the genetic organization of the ectoine/hydroxyectoine biosynthetic genes (Fig. 6) and those of proteins that are functionally associated with ectoine production, the specialized aspartokinase Ask_Ect [22,45] and the transcriptional regulator EctR [25] (Fig. 1). By analyzing the occurrence of these proteins on a genome-wide scale and by viewing them in a taxonomic context, we derived the currently most comprehensive *in silico* analysis of the production potential for the stress protectants and chemical chaperones ectoine and 5-hydroxyectoine in microorganisms (Fig. 1 and 6; Fig. S8). This dataset can therefore serve as a solid benchmark for future assessments as microbial genome and metagenomic sequence analysis continues to progress in a rapid pace.

Our analysis of the genetic organization of the *ectABC/D* biosynthetic genes revealed a robust arrangement into an operon-like structure in taxonomically widely separated microorganisms (Fig. 6). This assessment does not only provide clues for their potential transcriptional organization, but also gives hints about which of these gene clusters might be useful as building blocks for synthetic ectoine/hydroxyectoine production in heterologous host systems [22,56,71–74]. For instance, we surmise that the ectoine/hydroxyectoine biosynthetic genes from *Kytococcus sedentarius* [75] might be effectively exploited as a synthetic "bio-brick" for this purpose. In this microorganism, the genes for both enzymes (Ask_Ect and Asd) required for the synthesis of the direct ectoine precursor, L-aspartate-β-semialdehyde [17,18,65], seem to be co-transcribed with the *ectABCD* operon (Fig. 6B). Co-expression of the *ask_ect-asd-ectABCD* gene cluster should help to avoid the build-up of potential bottlenecks during heterologous ectoine/hydroxyectoine production for biotechnological and medical purposes [15,32,33].

We placed special emphasis in our study on the further biochemical [20,29] and structural analysis [44] of the ectoine hydroxylase. In terms of the EctD crystal structure, our new data reveal that the apo- and iron-liganded forms are virtually identical (Fig. 5A). Hence, the ectoine hydroxylase is pre-set in a configuration ready to accept the iron molecule (Fig. 5B) and the binding of the iron catalyst does not trigger large conformational changes. Together with the EctD proteins from *V. salexigens* and *S. coelicolor* that were previously studied biochemically [20,29], the six ectoine hydroxylases examined here define the salient biochemical features (Table 1 and 2) of this group of closely related enzymes (Fig. S8). The ectoine hydroxylases analyzed so far all possess similar kinetic parameters and catalytic efficiencies (Table 1) [20,29] but differ in their tolerance towards high temperature (Fig. 4) and in the influence of various salts on their enzyme activity (Table 2). It is hoped that the properties of some of the newly characterized EctD proteins will be suitable for further crystallographic studies so that a crystal structure of the ectoine hydroxylase with all its ligands (or its reaction product 5-hydroxyectoine) can be obtained in the future.

Materials and Methods

Chemicals

Ectoine and 5-hydroxyectoine were kind gifts from Dr. Thomas Lentzen and Dr. Irina Bagyan (bitop AG, Witten, Germany). 2-oxoglutarate (disodium salt) was obtained from Sigma-Aldrich (St. Louis, MO, USA). Anhydrotetracycline-hydrochloride (AHT), desthiobiotin, and Strep-Tactin Superflow chromatography material were purchased from IBA GmbH (Göttingen, Germany). X-Gal was obtained from AppliChem (Darmstadt, Germany), and the antibiotics kanamycin and ampicillin were purchased from Serva Electrophoresis GmbH (Heidelberg, Germany) and Carl Roth GmbH (Karlsruhe, Germany).

Bacteria, Media and Growth Conditions

The *Escherichia coli* strain DH5α (Invitrogen, Karlsruhe, Germany) was used as host for recombinant plasmids and as overproduction strain for EctD-proteins; it was maintained routinely on LB agar plates and liquid media [76]. When it contained recombinant plasmids, either ampicillin (100 μg ml^{-1}) or kanamycin (50 μg ml^{-1}) was added to the growth medium to select for the presence of the plasmids. When appropriate, X-gal was included in agar plates to screen for the insertion of the desired DNA fragments into the cloning vector pENTRY-IBA20 (IBA, Göttingen, Gemany). For the overproduction of EctD-type proteins, minimal medium A (MMA) [76] was used that was supplemented with 0.5% (w/v) glucose as the carbon source, 0.5% (w/v) casaminoacids, 1 mM $MgSO_4$, and 3 mM thiamine.

Recombinant DNA Techniques and Construction of Plasmids

All recombinant DNA techniques followed routine procedures. To construct expression plasmids carrying either the *H. elongata* or the *S. alaskensis ectD* gene with a C-terminal *Strep*-tag-II affinity peptide, we amplified these *ectD* genes from chromosomal DNA with PCR using custom synthesized DNA primers. A BsaI restriction site was introduced at both ends of the amplified DNA fragments allowing the directed insertion of the PCR products into the expression vector pASK-IBA3 (IBA, Göttingen, Gemany) via BsaI restriction and ligation reactions. The generated plasmids were pMP32 (*ectD* from *H. elongata*) and pMP40 (*ectD* from *S. alaskensis*). Expression plasmids carrying the *P. stutzeri*, *P. lautus*, *A. ehrlichii* or *A. cryptum ectD* gene with a C-terminal *Strep*-tag-II affinity peptide were constructed using the IBA Stargate cloning system (IBA, Göttingen, Gemany). The *ectD* gene from *P. stutzeri* was amplified from chromosomal DNA via PCR using custom synthesized primers that carried synthetically added LguI DNA restriction sites at their ends; this PCR fragment was cloned into the donor vector pENTRY-IBA20 via LguI restriction and concurrent ligation thereby yielding plasmid pMP34. DNA sequences from *P. lautus*, *A. ehrlichii* and *A. cryptum* genes were retrieved from the database and this information was used for codon-optimized synthesis of *ectD* genes (GeneScript, Piscataway, USA). An LguI restriction site was added to both ends of these genes, and they were inserted into the pENTRY-IBA20 donor

vector via LguI restriction and concurrent ligation. This generated plasmids pMP36 (*ectD* from *P. lautus*), pMP37 (*ectD* from *A. ehrlichii*), and pMP38 (*ectD* from *A. cryptum*). The synthetically manufactured *ectD* genes optimized for the expression in *E.coli* by GeneScript were deposited into the NCBI database with accession numbers JN019032 (*P. lautus ectD*), JN019031 (*A. ehrlichii ectD*) and JN019030 (*A. cryptum ectD*), respectively. To clone the *ectD* genes present on pMP34, pMP36, pMP37 and pMP38 into the pASG-IBA3 expression vector, Esp3I restriction and concurrent ligation of these plasmids and the expression vector pASG-IBA3 were carried out. In this way, in each of the recombinant *ectD* genes a short DNA fragment encoding the *Strep*-tag-II affinity peptide was added at their 3′-ends. The resulting plasmids were pMP41 (*ectD* from *P. stutzeri*), pMP43 (*ectD* from *P. lautus*), pMP44 (*ectD* from *A. ehrlichii*) and pMP48 (*ectD* from *A. cryptum*). The correct nucleotide sequence of all constructed plasmids was ascertained by DNA sequence analysis, which was carried out by Eurofins MWG Operon (Ebersberg, Germany).

Overproduction and Purification of Recombinant EctD Enzymes

In each of the constructed recombinant plasmids, the *ectD* gene is expressed from the *tet* promoter under the control of the AHT inducible TetR repressor (encoded by the *tetR* gene present on the expression vector). Overproduction of the different ectoine hydroxylases was performed in a chemically defined medium containing glucose as the carbon source essentially as previously described [43,44]. Briefly, cells of the *E. coli* strain DH5α harboring an appropriate plasmid were grown to an OD_{578} of about 0.7 at 37°C, the inducer AHT was then added to the culture to a final concentration of 0.2 mg mL^{-1}, and the growth temperature was then reduced to 35°C; growth of the cultures was continued for two hours. The cells were harvested by centrifugation (10 min, 5000 rpm, 4°C) and stored at −20°C until further used. A Strep-Tactin Superflow column was used to purify the recombinant EctD enzymes by affinity chromatography as detailed previously [20,43,44]. The purified EctD proteins were shock-frozen in liquid nitrogen and stored at −80°C until they were further used in HPLC-based enzyme activity assays. These EctD preparations typically contained between 0.87 and 0.96 mole iron per mol of EctD protein.

To provide large amounts of EctD protein for the crystallization trials, the above described overexpression protocol was varied somewhat. Cells of *E. coli* DH5α harboring a recombinant plasmid carrying an *ectD* gene were grown at 37°C to an OD_{578} of about 0.5 in a flask set on an aerial shaker (180 rpm). The cultivation temperature was then reduced to 30°C, and the shaker speed were decreased to 100 rpm. The cells were then grown to an OD_{578} of about 0.7, after which the inducer (AHT) of the TetR repressor was added to the cultures at a final concentration of 0.2 mg mL^{-1}. Cultures were grown for additional 2 hours and then harvested by centrifugation. By this modified overexpression protocol, the amount of purified EctD protein was increased from an average of 20–25 mg L^{-1} obtained by the initial protocol to 200–300 mg L^{-1}. The purity of the EctD was assessed by SDS-PAGE (12% polyacrylamide) and concentrated by ultra-filtration on spin columns (Sartorius Stedim Biotech GmBH, Göttingen, Germany) to about 10 mg ml^{-1} prior to the crystallization experiments.

Gel filtration chromatography was performed to determine the size of each individual purified EctD protein by loading 1 ml of each protein solution [5 mg ml^{-1}] onto a HiLoad 16/600 Superdex 200 pg column (GE Healthcare Europe GmbH, Freiburg, Germany) connected to an ÄKTA pure 25 L system (GE Healthcare Europe GmbH, Freiburg, Germany). The column was equilibrated and run in a 20 mM TES-buffer containing 150 NaCl. The evaluation of the column run was carried out with the Unicorn 6.3 software package (GE Healthcare Europe GmbH, Freiburg, Germany). A protein solution [3 mg ml^{-1}] of carbonic anhydrase (from bovine erythrocytes) (29 kDa), albumin (from bovine serum) (66 kDa), and alcohol dehydrogenase (from *Saccharomyces cerevisiae*) (150 kDa) was used as standard (Gel Filtration Markers Kit; Sigma-Aldrich, St. Louis, MO, USA).

Ectoine Hydroxylase Activity Assays

After affinity purification on Strep-Tactin Superflow material, the iron contents of the recombinantly produced EctD proteins were determined as described [77]. The hydroxylation of ectoine by EctD-type enzymes was measured by an HPLC-based enzyme assay [20]. In general, 30-µl reaction volumes containing 10 mM TES (pH 7.5), 1 mM $FeSO_4$, 10 mM 2-oxoglutarate, 6 mM ectoine, and various amounts of the purified EctD enzymes were incubated aerobically in an Eppendorf thermo-mixer (Hamburg, Germany) (set to 700 rpm) at 32°C for 20 min. The enzyme reaction was stopped by adding 30-µl acetonitrile (100%) to the reaction mixture, immediately followed by centrifugation (10 min, 4°C, 32000×*g*) to remove the denatured EctD protein. The conversion of ectoine to 5-hydroxyectoine was assessed by loading 20 µl of the reaction mixture supernatant onto a GROM-SIL Amino-1PR column (125 mm by 4 mm; 3 µm particle size (GROM, Rottenburg-Hailfingen, Germany) attached to a UV-visible detector system (LINEAR UVIS 205; SYKAM, Fürstenfeldbruck, Germany) in an HPLC system (SYKAM). The absorbance of ectoine and 5-hydroxyectoine was monitored at 210 nm [11,20], and the amount of 5-hydroxyectoine formed was quantitated using the ChromStar 7.0 software package (SYKAM, Fürstenfeldbruck, Germany).

To determine the biochemical properties of the different EctD-type proteins, the above-described standard enzyme assay was modified with respect to the incubation temperature, the buffer and pH conditions, and the salt content of the assay solution. To determine the kinetic parameters of the studied ectoine hydroxylases, each of the different EctD enzymes was assayed at its optimal conditions (Table 1) with varied concentrations of either ectoine (between 0 and 80 mM) or 2-oxoglutarate (between 0 and 50 mM). To assess the ability of the EctD protein to perform the reverse enzyme reaction (forming ectoine from 5-hydroxyectoine), samples of the purified *V. salexigens* EctD protein were incubated under the assay conditions described above, except that various concentrations of 5-hydroxyectoine (from 6 mM to 100 mM) instead of ectoine, succinate (from 5 to 40 mM) instead of 2-oxoglutarate, and bicarbonate (between 5 mM to 20 mM) were used. These reaction samples were incubated (either with or without shaking in a thermo-mixer) for various time periods (from 20 min to 24 hours), and processed as described above. The products of the enzyme reactions were then analyzed by HPLC [20].

Database Searches, Alignments of Amino Acid Sequences, and Construction of Phylogenetic Trees of EctC- and EctD-type Proteins

The bioinformatics tools available at the DOE Joint Genome Institute website (http://www.jgi.doe.gov) [46] were used to retrieve EctC- and EctD-type protein sequences from finished microbial genomes (search date: 07/31/2013). For these database searches, the amino acid sequences of the EctC (accession number: AAY29688) and EctD (accession number: AAY29689) proteins from *V. salexigens* [20] were used as the query sequence using the

BLAST program [78]. The retrieved EctC and EctD protein sequences were aligned and compared using ClustalW [62]. Based on these alignments, phylogenetic trees were calculated using the iTOL-software package (http://itol.embl.de/) [61] to visualize the distribution of EctC and EctD proteins among members of the *Bacteria* and *Archaea*. The genetic organization of the *ectABC*/(*ectD*) gene cluster and its flanking sequences were analyzed using the online tool available from the DOE Joint Genome Institute website [46].

Crystallization of the *V. Salexigens* EctD Protein in its Iron-free Form

Crystallization trials were performed using the sitting-drop vapor diffusion method at 20°C. A homogenous protein solution of the affinity-purified EctD protein (in 20 mM TES pH 7.5, 80 mM NaCl) was concentrated to 10 mg/ml prior to crystallization experiments. EctD crystals were grown by mixing 1.5 µl protein solution with 1.5 µl reservoir solution containing 100 mM MES pH 5.0 and 1.2 M ammonium sulfate; the EctD crystals grew within 6–12 days to their final size of around $80\times90\times100$ µm^3. Crystals were cryoprotected by carefully adding 1 µl 100% glycerol to the crystallization drop before freezing the crystals in liquid nitrogen.

Data Collection, Refinement and Crystallographic Analysis of the EctD Protein

EctD crystals diffracted X-rays to a minimum resolution of 1.85 Å for the apo-EctD. The dataset was collected at the ID23-EH2 beamline at the ESRF (Grenoble, France) and processed with XDS [79]. The crystal structure of the iron-bound *V. salexigens* EctD protein (PDB code: 3EMR) [44] was used as a template to obtain initial phases using PHASER [80]. The structure was further refined using REFMAC5 [81] and manually adjusted using COOT [82]. Dataset and refinement statistics for the apo-EctD crystal structure are listed in Table S1 and were analyzed with Procheck [83]. The crystallographic information for the *V. salexigens* apo-EctD protein was deposited in the Protein Data Base (PDB) [84] with the PDB accession code 4NMI. Figures of protein molecules derived from crystal structures were prepared using the PyMol software suit (www.pymol.org).

Supporting Information

Figure S1 Biochemical properties of the EctD enzyme from *Halomonas elongata*. The enzyme activity of the ectoine hydroxylase from *H. elongata* is shown with respect to (A) the temperature optimum, (B) the pH optimum, and the influence of different salts: (C) potassium chloride, (D) sodium chloride, (E) potassium glutamate, and (F) ammonium chloride.
(TIF)

Figure S2 Biochemical properties of the EctD enzyme from *Pseudomonas stutzeri*. The enzyme activity of the ectoine hydroxylase from *P. stutzeri* is shown with respect to (A) the temperature optimum, (B) the pH optimum, and the influence of different salts: (C) potassium chloride, (D) sodium chloride, (E) potassium glutamate, and (F) ammonium chloride.
(TIF)

Figure S3 Biochemical properties of the EctD enzyme from *Paenibacillus lautus*. The enzyme activity of the ectoine hydroxylase from *P. lautus* is shown with respect to (A) the temperature optimum, (B) the pH optimum, and the influence of different salts: (C) potassium chloride, (D) sodium chloride, (E) potassium glutamate, and (F) ammonium chloride.
(TIF)

Figure S4 Biochemical properties of the EctD enzyme from *Alkalilimnicola ehrlichii*. The enzyme activity of the ectoine hydroxylase from *A. ehrlichii* is shown with respect to (A) the temperature optimum, (B) the pH optimum, and the influence of different salts: (C) potassium chloride, (D) sodium chloride, (E) potassium glutamate, and (F) ammonium chloride.
(TIF)

Figure S5 Biochemical properties of the EctD enzyme from *Acidiphilium cryptum*. The enzyme activity of the ectoine hydroxylase from *A. cryptum* is shown with respect to (A) the temperature optimum, (B) the pH optimum, and the influence of different salts: (C) potassium chloride, (D) sodium chloride, (E) potassium glutamate, and (F) ammonium chloride.
(TIF)

Figure S6 Gel filtration analysis of the *Sphingopyxis alaskensis* EctD protein. The *S. alaskensis* EctD protein was purified by affinity chromatography and its quaternary structure was then assessed by gel filtration analysis on a HiLoad 16/600 Superdex 200 pg column. The column was equilibrated and run in a 20 mM TES-buffer containing 150 NaCl. A protein solution [3 mg/ml] of carbonic anhydrase (from bovine erythrocytes) (29 kDa), albumin (from bovine serum) (66 kDa), and alcohol dehydrogenase (from *Saccharomyces cerevisiae*) (150 kDa) was used as a standard. The calculated molecular mass of the *S. alaskensis* EctD protein with the attached *Strep*-tag-II affinity peptide (nine amino acids) is 35.29 kDa; the molecular mass calculated from the column run was 70.38 kDa. Arrows indicate the elution of the standard proteins from the gel filtration column. mAU: milli absorption units.
(TIF)

Figure S7 Enzyme activity of the ectoine hydroxylase is not reversible. The forward and backward enzyme reactions of the EctD protein from *S. alaskensis* were tested, and the formation of ectoine and hydroxyectoine was monitored by HPLC analysis. (a) Chromatograms from HPLC measurements monitored at 210 nm of a mixture of commercially available ectoine and 5-hydroxyectoine standards. (b) HPLC tracing of the EctD-catalyzed enzyme reaction mixture that initially contained 6 mM ectoine; the enzyme assay was run for 20 min. (c) HPLC tracing of the EctD-catalyzed "reverse" enzyme reaction mixture that initially contained 6 mM hydroxyectoine; the enzyme assay was run for 24 h.
(TIF)

Figure S8 Phylogenetic tree of EctD-type proteins. The phylogenetic tree of ectoine hydroxylases shown is based on the alignment of EctD amino acid sequences identified by a BLAST search at the JGI Web-server, and that were then aligned using ClustalW. The phylogenetic distribution of the aligned EctD proteins was assessed via the iTOL Web-server. Evolutionary distances are not given. The color code indicates the distribution of EctD among members of the *Bacteria* and *Archaea*.
(TIF)

Table S1 Data collection and refinement statistics for the crystal structure of the EctD protein from *V. salexigens* in its iron-free form.
(DOC)

Acknowledgments

We thank Jochen Sohn for excellent technical assistance and Georg Lentzen and Irina Bagyan (Bitop AG, Witten, Germany) for their kind gifts of ectoine and 5-hydroxyectoine. We greatly appreciate the expert help of

Vickie Koogle in the language editing of our manuscript and thank Lutz Schmitt for helpful discussions. E.B. greatly valued the hospitality and kind support of Tom Silhavy during a sabbatical at the Department of Molecular Biology of Princeton University (Princeton, NJ, USA). N.W. and E.B. are very grateful to Rolf Thauer for his continued support.

Author Contributions

Conceived and designed the experiments: EB SHJS. Performed the experiments: NW AH MP. Analyzed the data: NW AH MP JH SHJS EB. Wrote the paper: SHJS JH EB.

References

1. Bremer E, Krämer R (2000) Coping with osmotic challenges: osmoregulation through accumulation and release of compatible solutes. In: Storz G, Hengge-Aronis R, editors. Bacterial Stress Responses: ASM Press. 79 97.
2. Wood JM, Bremer E, Csonka LN, Krämer R, Poolman B, et al. (2001) Osmosensing and osmoregulatory compatible solute accumulation by bacteria. Comp Biochem Physiol Part A Mol Integr Physiol 130: 437 460.
3. da Costa MS, Santos H, Galinski EA (1998) An overview of the role and diversity of compatible solutes in Bacteria and Archaea. Adv Biochem Eng Biotechnol 61: 117 153.
4. Kempf B, Bremer E (1998) Uptake and synthesis of compatible solutes as microbial stress responses to high osmolality environments. Arch Microbiol 170: 319 330.
5. Roeβler M, Müller V (2001) Osmoadaptation in bacteria and archaea: common principles and differences. Env Microbiol Rep 3: 743 754.
6. Wood JM (2011) Bacterial osmoregulation: a paradigm for the study of cellular homeostasis. Annu Rev Microbiol 65: 215 238.
7. Street TO, Bolen DW, Rose GD (2006) A molecular mechanism for osmolyte-induced protein stability. Proc Natl Acad Sci USA 103: 13997 14002.
8. Cayley S, Record MT Jr (2003) Roles of cytoplasmic osmolytes, water, and crowding in the response of *Escherichia coli* to osmotic stress: biophysical basis of osmoprotection by glycine betaine. Biochemistry 42: 12596 12609.
9. Brown AD (1976) Microbial water stress. Bacteriol Rev 40: 803 846.
10. Hoffmann T, Wensing A, Brosius M, Steil L, Völker U, et al. (2013) Osmotic control of *opuA* expression in *Bacillus subtilis* and its modulation in response to intracellular glycine betaine and proline pools. J Bacteriol 195: 510 522.
11. Kuhlmann AU, Bremer E (2002) Osmotically regulated synthesis of the compatible solute ectoine in *Bacillus pasteurii* and related Bacillus spp. Appl Env Microbiol 68: 772 783.
12. Cayley S, Lewis BA, Record MT Jr (1992) Origins of the osmoprotective properties of betaine and proline in *Escherichia coli* K-12. J Bacteriol 174: 1586 1595.
13. Inbar L, Lapidot A (1988) The structure and biosynthesis of new tetrahydropyrimidine derivatives in actinomycin D producer *Streptomyces parvulus*. Use of ^{13}C- and ^{15}N-labeled L-glutamate and ^{13}C and ^{15}N NMR spectroscopy. J Biol Chem 263: 16014 16022.
14. Galinski EA, Pfeiffer HP, Trüper HG (1985) 1,4,5,6-Tetrahydro-2-methyl-4-pyrimidinecarboxylic acid. A novel cyclic amino acid from halophilic phototrophic bacteria of the genus *Ectothiorhodospira*. Eur J Biochem 149: 135 139.
15. Pastor JM, Salvador M, Argandona M, Bernal V, Reina-Bueno M, et al. (2010) Ectoines in cell stress protection: uses and biotechnological production. Biotechnol Adv 28: 782 801.
16. Galinski EA, Trüper HG (1994) Microbial behaviour in salt-stressed ecosystems. FEMS Microbiol Rev 15: 95 108.
17. Ono H, Sawada K, Khunajakr N, Tao T, Yamamoto M, et al. (1999) Characterization of biosynthetic enzymes for ectoine as a compatible solute in a moderately halophilic eubacterium, *Halomonas elongata*. J Bacteriol 181: 91 99.
18. Peters P, Galinski EA, Trüper HG (1990) The biosyntheis of ectoine. FEMS Microbiol Lett 71: 157 162.
19. Louis P, Galinski EA (1997) Characterization of genes for the biosynthesis of the compatible solute ectoine from *Marinococcus halophilus* and osmoregulated expression in *Escherichia coli*. Microbiology 143: 1141 1149.
20. Bursy J, Pierik AJ, Pica N, Bremer E (2007) Osmotically induced synthesis of the compatible solute hydroxyectoine is mediated by an evolutionarily conserved ectoine hydroxylase. J Biol Chem 282: 31147 31155.
21. Calderon MI, Vargas C, Rojo F, Iglesias-Guerra F, Csonka LN, et al. (2004) Complex regulation of the synthesis of the compatible solute ectoine in the halophilic bacterium *Chromohalobacter salexigens* DSM 3043^{T}. Microbiology 150: 3051 3063.
22. Stöveken N, Pittelkow M, Sinner T, Jensen RA, Heider J, et al. (2011) A specialized aspartokinase enhances the biosynthesis of the osmoprotectants ectoine and hydroxyectoine in *Pseudomonas stutzeri* A1501. J Bacteriol 193: 4456 4468.
23. Saum SH, Müller V (2008) Regulation of osmoadaptation in the moderate halophile *Halobacillus halophilus*: chloride, glutamate and switching osmolyte strategies. Saline Systems 4: 4.
24. Mustakhimov II, Reshetnikov AS, Fedorov DN, Khmelenina VN, Trotsenko YA (2012) Role of EctR as transcriptional regulator of ectoine biosynthesis genes in *Methylophaga thalassica*. Biochemistry (Mosc) 77: 857 863.
25. Mustakhimov II, Reshetnikov AS, Glukhov AS, Khmelenina VN, Kalyuzhnaya MG, et al. (2010) Identification and characterization of EctR1, a new transcriptional regulator of the ectoine biosynthesis genes in the halotolerant methanotroph *Methylomicrobium alcaliphilum* 20Z. J Bacteriol 192: 410 417.
26. Kuhlmann AU, Bursy J, Gimpel S, Hoffmann T, Bremer E (2008) Synthesis of the compatible solute ectoine in *Virgibacillus pantothenticus* is triggered by high salinity and low growth temperature. Appl Environ Microbiol 74: 4560 4563.
27. Garcia-Estepa R, Argandona M, Reina-Bueno M, Capote N, Iglesias-Guerra F, et al. (2006) The *ectD* gene, which is involved in the synthesis of the compatible solute hydroxyectoine, is essential for thermoprotection of the halophilic bacterium *Chromohalobacter salexigens*. J Bacteriol 188: 3774 3784.
28. Kuhlmann AU, Hoffmann T, Bursy J, Jebbar M, Bremer E (2011) Ectoine and hydroxyectoine as protectants against osmotic and cold stress: uptake through the SigB-controlled betaine-choline- carnitine transporter-type carrier EctT from *Virgibacillus pantothenticus*. J Bacteriol 193: 4699 4708.
29. Bursy J, Kuhlmann AU, Pittelkow M, Hartmann H, Jebbar M, et al. (2008) Synthesis and uptake of the compatible solutes ectoine and 5-hydroxyectoine by *Streptomyces coelicolor* A3(2) in response to salt and heat stresses. Appl Environ Microbiol 74: 7286 7296.
30. Inbar L, Frolow F, Lapidot A (1993) The conformation of new tetrahydropyrimidine derivatives in solution and in the crystal. Eur J Biochem 214: 897 906.
31. Prabhu J, Schauwecker F, Grammel N, Keller U, Bernhard M (2004) Functional expression of the ectoine hydroxylase gene (*thpD*) from *Streptomyces chrysomallus* in *Halomonas elongata*. Appl Environ Microbiol 74: 3130 3132.
32. Lentzen G, Schwarz T (2006) Extremolytes: natural compounds from extremophiles for versatile applications. Appl Microbiol Biotechnol 72: 623 634.
33. Graf R, Anzali S, Buenger J, Pfluecker F, Driller H (2008) The multifunctional role of ectoine as a natural cell protectant. Clin Dermatol 26: 326 333.
34. Schwibbert K, Marin-Sanguino A, Bagyan I, Heidrich G, Lentzen G, et al. (2011) A blueprint of ectoine metabolism from the genome of the industrial producer *Halomonas elongata* DSM 2581 T. Environ Microbiol 13: 1973 1994.
35. Borges N, Ramos A, Raven ND, Sharp RJ, Santos H (2002) Comparative study of the thermostabilizing properties of mannosylglycerate and other compatible solutes on model enzymes. Extremophiles 6: 209 216.
36. Lippert K, Galinski EA (1992) Enzyme stabilization by ectoine-type compatible solutes: protection against heating, freezing and drying. Appl Microbial Biotechnol 37: 61 65.
37. Kurz M (2008) Compatible solute influence on nucleic acids: many questions but few answers. Saline Systems 4: 6.
38. Van-Thuoc D, Hashim SO, Hatti-Kaul R, Mamo G (2013) Ectoine-mediated protection of enzyme from the effect of pH and temperature stress: a study using *Bacillus halodurans* xylanase as a model. Appl Microbiol Biotechnol 97: 6271 6278.
39. Hausinger RP (2004) FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes. Crit Rev Biochem Mol Biol 39: 21 68.
40. Aik W, McDonough MA, Thalhammer A, Chowdhury R, Schofield CJ (2012) Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. Cur Opin Struct Biol 22: 691 700.
41. Hangasky JA, Taabazuing CY, Valliere MA, Knapp MJ (2013) Imposing function down a (cupin)-barrel: secondary structure and metal stereochemistry in the alphaKG-dependent oxygenases. Metallomics 5: 287 301.
42. Grzyska PK, Appelman EH, Hausinger RP, Proshlyakov DA (2010) Insight into the mechanism of an iron dioxygenase by resolution of steps following the FeIV = HO species. Proc Natl Acad Sci USA 107: 3982 3987.
43. Widderich N, Pittelkow M, Höppner A, Mulnaes D, Buckel W, et al. (2013) Molecular dynamics simulations and structure-guided mutagenesis provide insight into the architecture of the catalytic core of the ectoine hydroxylase. J Mol Biol 426: 586 600.
44. Reuter K, Pittelkow M, Bursy J, Heine A, Craan T, et al. (2010) Synthesis of 5-hydroxyectoine from ectoine: crystal structure of the non-heme iron(II) and 2-oxoglutarate-dependent dioxygenase EctD. PloS one 5: e10647.
45. Reshetnikov AS, Khmelenina VN, Trotsenko YA (2006) Characterization of the ectoine biosynthesis genes of haloalkalotolerant obligate methanotroph "*Methylomicrobium alcaliphilum* 20Z". Arch Microbiol 184: 286 297.
46. Nordberg H, Cantor M, Dusheyko S, Hua S, Poliakov A, et al. (2013) The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic Acids Res 42: D26 D31.
47. Polz MF, Alm EJ, Hanage WP (2013) Horizontal gene transfer and the evolution of bacterial and archaeal population structure. Trends Genet 29: 170 175.
48. Vreeland RH, Lichtfiled CD, Martin EL, Elliot E (1980) *Halomonas elongata*, a new genus and species of extremely salt-tolerant bacteria. Int J Syst Bacteriol 30: 485 495.
49. Küsel K, Dorsch T, Acker G, Stackebrandt E (1999) Microbial reduction of Fe(III) in acidic sediments: isolation of *Acidiphilium cryptum* JF-5 capable of coupling the reduction of Fe(III) to the oxidation of glucose. Appl Environ Microbiol 65: 3633 3640.
50. Bilgin AA, Silverstein J, Jenkins JD (2004) Iron respiration by *Acidiphilium cryptum* at pH 5. FEMS Microbiol Ecol 49: 137 143.

51. Hoeft SE, Blum JS, Stolz JF, Tabita FR, Witte B, et al. (2007) *Alkalilimnicola ehrlichii* sp. nov., a novel, arsenite-oxidizing haloalkaliphilic gammaproteobacterium capable of chemoautotrophic or heterotrophic growth with nitrate or oxygen as the electron acceptor. Int J Syst Evol Microbiol 57: 504 512.
52. Ting L, Williams TJ, Cowley MJ, Lauro FM, Guilhaus M, et al. (2010) Cold adaptation in the marine bacterium, *Sphingopyxis alaskensis*, assessed using quantitative proteomics. Environ Microbiol 12: 2658 2676.
53. Vancanneyt M, Schut F, Snauwaert C, Goris J, Swings J, et al. (2001) *Sphingomonas alaskensis* sp. nov., a dominant bacterium from a marine oligotrophic environment. Int J Syst Evol Microbiol 51: 73 79.
54. Mead DA, Lucas S, Copeland A, Lapidus A, Cheng JF, et al. (2012) Complete genome sequence of *Paenibacillus* strain Y4.12MC10, a novel *Paenibacillus lautus* strain isolated from Obsidian Hot Spring in Yellowstone National Park. Stand Genomic Sci 6: 381 400.
55. Yan Y, Yang J, Dou Y, Chen M, Ping S, et al. (2008) Nitrogen fixation island and rhizosphere competence traits in the genome of root-associated *Pseudomonas stutzeri* A1501. Proc Natl Acad Sci USA 105: 7564 7569.
56. Seip B, Galinski EA, Kurz M (2011) Natural and engineered hydroxyectoine production based on the *Pseudomonas stutzeri ectABCD-ask* gene cluster. Appl Environ Microbiol 77: 1368 1374.
57. Witt EM, Davies NW, Galinski EA (2011) Unexpected property of ectoine synthase and its application for synthesis of the engineered compatible solute ADPC. Appl Microbiol Biotechnol 91: 113 122.
58. Yan D, Lenz P, Hwa T (2011) Overcoming fluctuation and leakage problems in the quantification of intracellular 2-oxoglutarate levels in *Escherichia coli*. Appl Environ Microbiol 77: 6763 6771.
59. Straganz GD, Nidetzky B (2006) Variations of the 2-His-1-carboxylate theme in mononuclear non-heme FeII oxygenases. Chembiochem 10: 1536 1548.
60. Reshetnikov AS, Khmelenina VN, Mustakhimov, II, Kalyuzhnaya M, Lidstrom M, et al. (2011) Diversity and phylogeny of the ectoine biosynthesis genes in aerobic, moderately halophilic methylotrophic bacteria. Extremophiles 15: 653 663.
61. Letunic I, Bork P (2011) Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Res 39: W475 478.
62. Thompson JD, Plewniak F, Thierry JC, Poch O (2000) DbClustal: rapid and reliable global multiple alignments of protein sequences detected by database searches. Nucleic Acids Res 28: 2919 2926.
63. Kurz M, Burch AY, Seip B, Lindow SE, Gross H (2010) Genome-driven investigation of compatible solute biosynthesis pathways of *Pseudomonas syringae* pv. syringae and their contribution to water stress tolerance. Appl Environ Microbiol 76: 5452 5462.
64. Spang A, Hatzenpichler R, Brochier-Armanet C, Rattei T, Tischler P, et al. (2010) Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota. Trends Microbiol 18: 331 340.
65. Lo CC, Bonner CA, Xie G, D'Souza M, Jensen RA (2009) Cohesion group approach for evolutionary analysis of aspartokinase, an enzyme that feeds a branched network of many biochemical pathways. Microbiol Mol Biol Rev 73: 594 651.
66. Bestvater T, Louis P, Galinski EA (2008) Heterologous ectoine production in *Escherichia coli:* by-passing the metabolic bottle-neck. Saline Systems 4, 12.
67. Shikuma NJ, Davis KR, Fong JNC, Yildiz FH (2013) The transcriptional regulator, CosR, controls compatible solute biosynthesis and transport, motility and biofilm formation in *Vibrio cholerae*. Environ Microbiol 15: 1387 1399.
68. Pflughoeft KJ, Kierek K, Watnick PI (2003) Role of ectoine in *Vibrio cholerae* osmoadaptation. Appl Environ Microbiol 69: 5919 5927.
69. Ongagna-Yhombi SY, Boyd EF (2013) Biosynthesis of the osmoprotectant ectoine, but not glycine betaine, is critical for survival of osmotically stressed *Vibrio parahaemolyticus* cells. Appl Environ Microbiol 79: 5038 5049.
70. Vargas C, Argandona M, Reina-Bueno M, Rodriguez-Moya J, Fernandez-Aunion C, et al. (2008) Unravelling the adaptation responses to osmotic and temperature stress in *Chromohalobacter salexigens*, a bacterium with broad salinity tolerance. Saline Systems 4: 14.
71. Schubert T, Maskow T, Benndorf D, Harms H, Breuer U (2007) Continuous synthesis and excretion of the compatible solute ectoine by a transgenic, nonhalophilic bacterium. Appl Environ Microbiol 73: 3343 3347.
72. Rodriguez-Moya J, Argandona M, Iglesias-Guerra F, Nieto JJ, Vargas C (2013) Temperature- and salinity-decoupled overproduction of hydroxyectoine by *Chromohalobacter salexigens*. Appl Environ Microbiol 79: 1018 1023.
73. Eilert E, Kranz A, Hollenberg CP, Piontek M, Suckow M (2013) Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha*. J Biotechnol 167: 85 93.
74. Becker J, Schäfer R, Kohlstedt M, Harder BJ, Borchert NS, et al. (2013) Systems metabolic engineering of *Corynebacterium glutamicum* for production of the chemical chaperone ectoine. Microb Cell Fact 12: 110.
75. Sims D, Brettin T, Detter JC, Han C, Lapidus A, et al. (2009) Complete genome sequence of *Kytococcus sedentarius* type strain (541). Stand Genomic Sci 1: 12 20.
76. Miller JH (1972) Experiments in molecular genetics. Cold Spring Harbor New York: Cold Spring Harbor Laboratory.
77. Lovenberg W, Buchanan BB, Rabinowitz JC (1963) Studies on the chemical nature of Clostridial ferredoxin. J Biol Chem 238: 3899 3913.
78. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403 410.
79. Kabsch W (1993) Automatic processing of rotatation diffraction data from crystalls of initially unknown symetry and cell constants. J Appl Crystallogr 26: 795 800.
80. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, et al. (2007) Phaser crystallographic software. J Appl Crystallogr 40: 658 674.
81. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53: 240 255.
82. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60: 2126 2132.
83. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26: 283 291.
84. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The protein data bank. Nucleic Acids Res 28: 235 342.

SUPPLEMENTARY MATERIAL

Biochemical Properties of Ectoine Hydroxylases from Extremophiles and Their Wider Taxonomic Distribution among Microorganisms

Nils Widderich[1,2¶], Astrid Hoeppner[3,¶], Marco Pittelkow[1], Johann Heider[1,5], Sander H.J. Smits[4*], Erhard Bremer[1,5*]

[1] Laboratory for Microbiology, Department of Biology, Philipps-University Marburg, Marburg, Germany

[2] Max Planck Institute for Terrestrial Microbiology, Emeritus Group R.K. Thauer, Marburg, Germany

[3] X-Ray Facility and Crystal Farm, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

[4] Institute of Biochemistry, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

[5] LOEWE-Center for Synthetic Microbiology, Philipps-University Marburg, Marburg, Germany

Running title: Ectoine and Its Derivative 5-Hydroxyectoine

[¶] These authors contributed equally to this work

* Corresponding authors e-mail: sander.smits@hhu.de (SHGS); bremer@biologie.uni-marburg.de (EB)

Supplemental figures

Widderich *et al.*, Figure S1

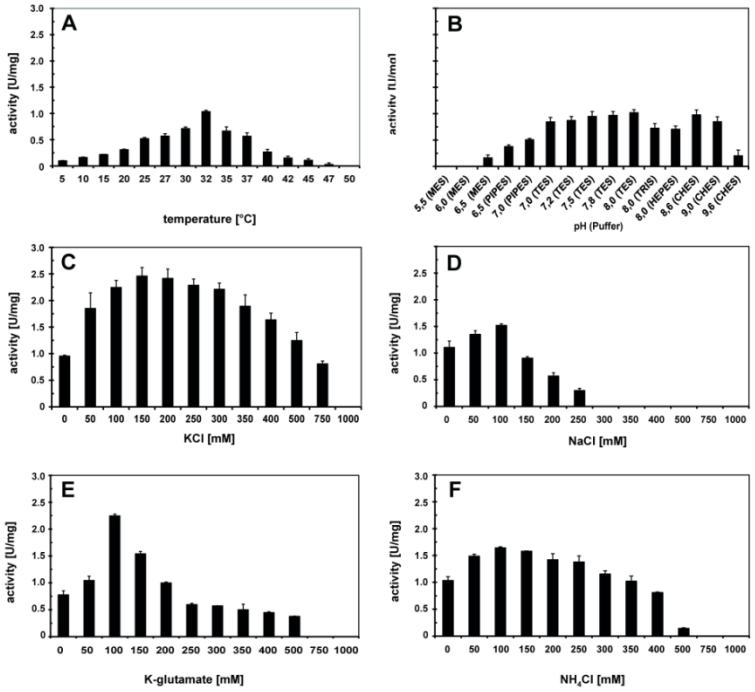


Figure S1 Biochemical properties of the EctD enzyme from *Halomonas elongata*. The enzyme activity of the ectoine hydroxylase from *H. elongata* is shown with respect to (A) the temperature optimum, (B) the pH optimum, and the influence of different salts: (C) potassium chloride, (D) sodium chloride, (E) potassium glutamate, and (F) ammonium chloride.

Supplemental figures

Widderich *et al.*, Figure S2

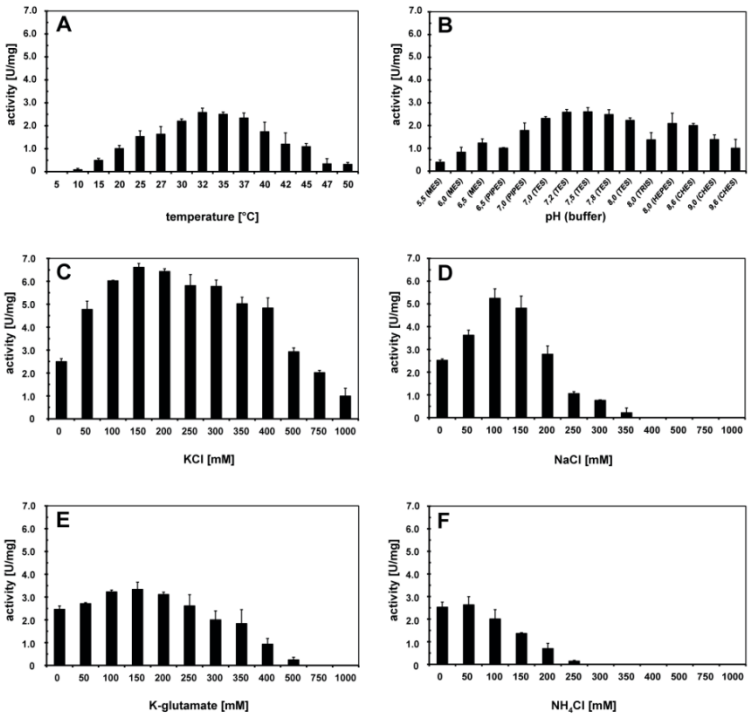


Figure S2 Biochemical properties of the EctD enzyme from *Pseudomonas stutzeri*. The enzyme activity of the ectoine hydroxylase from *P. stutzeri* is shown with respect to (A) the temperature optimum, (B) the pH optimum, and the influence of different salts: (C) potassium chloride, (D) sodium chloride, (E) potassium glutamate, and (F) ammonium chloride.

Supplemental figures

Widderich *et al.*, Figure S3

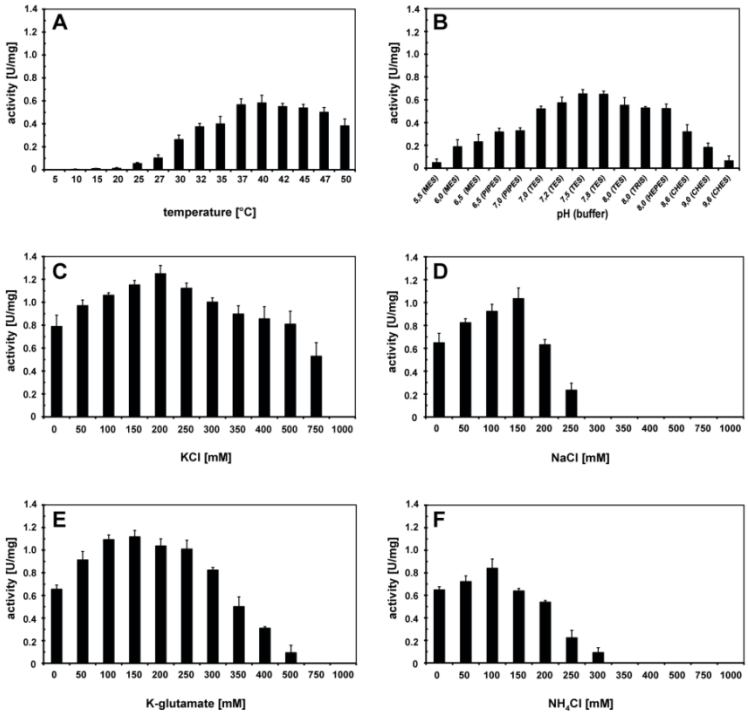


Figure S3 Biochemical properties of the EctD enzyme from *Paenibacillus lautus*. The enzyme activity of the ectoine hydroxylase from *P. lautus* is shown with respect to (A) the temperature optimum, (B) the pH optimum, and the influence of different salts: (C) potassium chloride, (D) sodium chloride, (E) potassium glutamate, and (F) ammonium chloride.

Supplemental figures

Widderich *et al.*, Figure S4

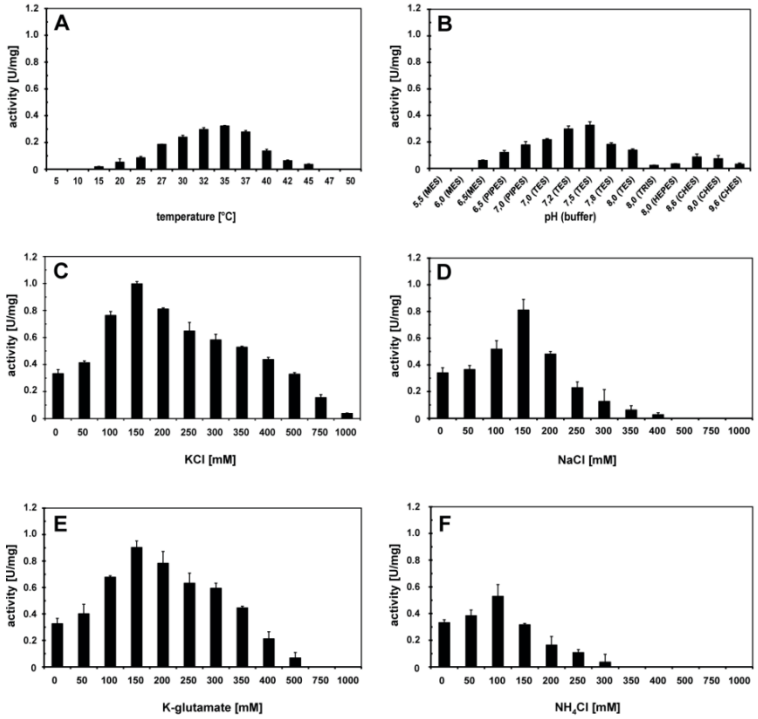


Figure S4 Biochemical properties of the EctD enzyme from *Alkalilimnicola ehrlichii*. The enzyme activity of the ectoine hydroxylase from *A. ehrlichii* is shown with respect to (A) the temperature optimum, (B) the pH optimum, and the influence of different salts: (C) potassium chloride, (D) sodium chloride, (E) potassium glutamate, and (F) ammonium chloride.

Supplemental figures

Widderich *et al.*, Figure S5

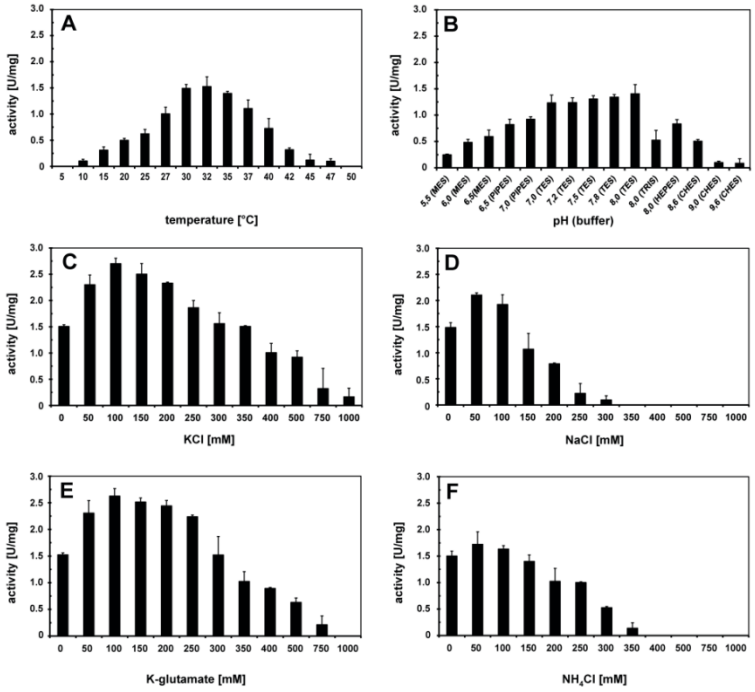


Figure S5 Biochemical properties of the EctD enzyme from *Acidiphilium cryptum*. The enzyme activity of the ectoine hydroxylase from *A. cryptum* is shown with respect to (A) the temperature optimum, (B) the pH optimum, and the influence of different salts: (C) potassium chloride, (D) sodium chloride, (E) potassium glutamate, and (F) ammonium chloride.

Supplemental figures

Widderich *et al.*, Figure S6

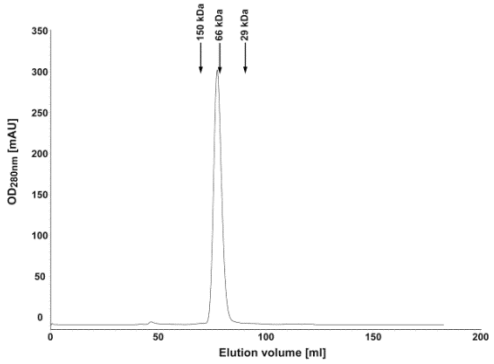


Figure S6 Gel filtration analysis of the *Sphingopyxis alaskensis* EctD protein. The *S. alaskensis* EctD protein was purified by affinity chromatography and its quaternary structure was then assessed by gel filtration analysis on a HiLoad 16/600 Superdex 200 pg column. The column was equilibrated and run in a 20 mM TES-buffer containing 150 NaCl. A protein solution [3 mg/ml] of carbonic anhydrase (from bovine erythrocytes) (29 kDa), albumin (from bovine serum) (66 kDa), and alcohol dehydrogenase (from Saccharomyces cerevisiae) (150 kDa) was used as a standard. The calculated molecular mass of the *S. alaskensis* EctD protein with the attached Strep-tag-II affinity peptide (nine amino acids) is 35.29 kDa; the molecular mass calculated from the column run was 70.38 kDa. Arrows indicate the elution of the standard proteins from the gel filtration column. mAU: milli absorption units.

Supplemental figures

Widderich *et al.*, Figure S7

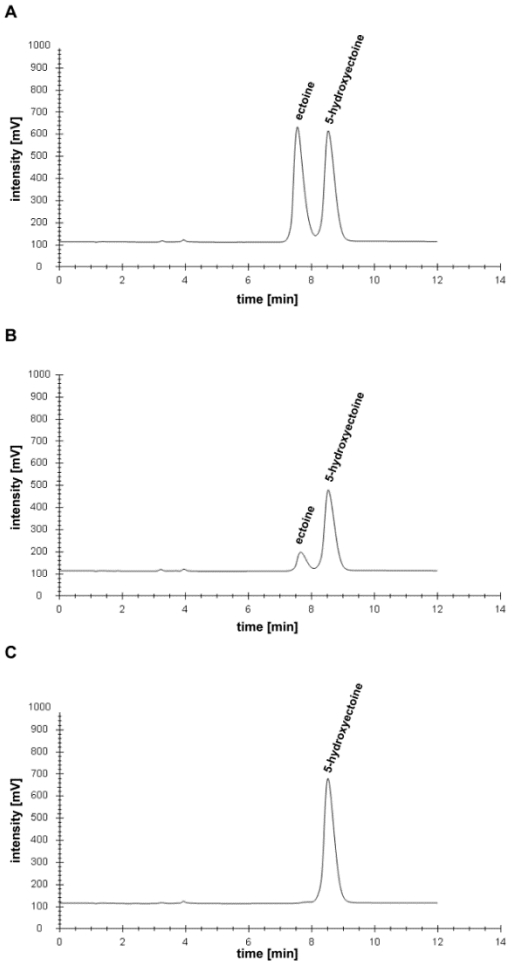


Figure S7 Enzyme activity of the ectoine hydroxylase is not reversible. The forward and backward enzyme reactions of the EctD protein from *S. alaskensis* were tested, and the formation of ectoine and hydroxyectoine was monitored by HPLC analysis. (a) Chromatograms from HPLC measurements monitored at 210 nm of a mixture of commercially available ectoine and 5-hydroxyectoine standards. (b) HPLC tracing of the EctD-catalyzed enzyme reaction mixture that initially contained 6 mM ectoine; the enzyme assay was run for 20 min. (c) HPLC tracing of the EctD-catalyzed ‘‘reverse’’ enzyme reaction mixture that initially contained 6 mM hydroxyectoine; the enzyme assay was run for 24 h.

Supplemental figures

Widderich *et al.*, Figure S8

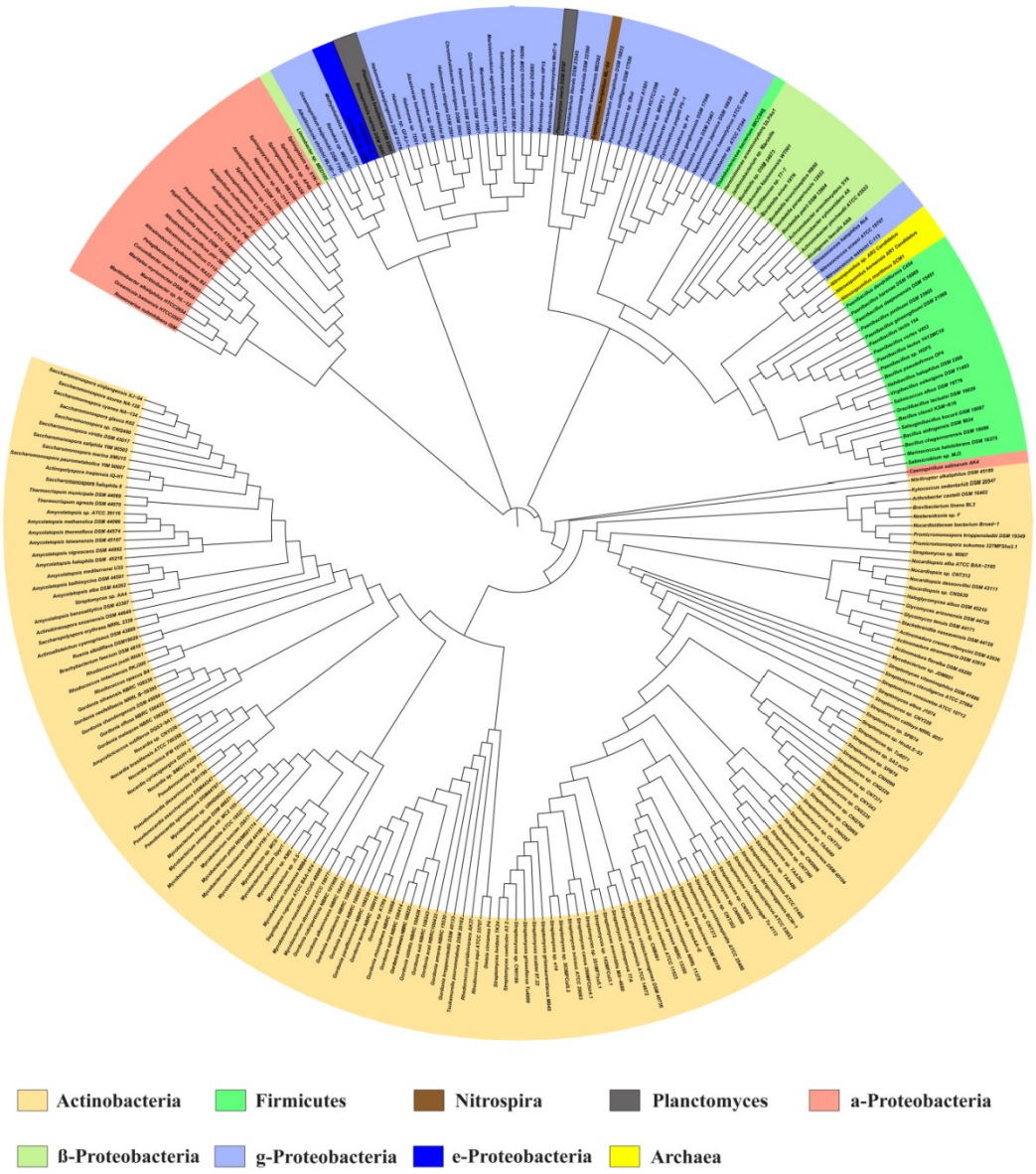


Figure S8 Phylogenetic tree of EctD-type proteins. The phylogenetic tree of ectoine hydroxylases shown is based on the alignment of EctD amino acid sequences identified by a BLAST search at the JGI Web-server, and that were then aligned using ClustalW. The phylogenetic distribution of the aligned EctD proteins was assessed via the iTOL Web-server. Evolutionary distances are not given. The color code indicates the distribution of EctD among members of the *Bacteria* and *Archaea*.

Supplemental tables

Widderich *et al.*, Table S1

Table S1

A. Crystal parameters at 100 K	
Space group	$P6_522$
Unit Cell parameters	
a, b, c (Å)	102.89 102.89 159.65
α=β(°)	90.0
χ (°)	120.0
B. Data collection and processing	
Wavelength (Å)	0.87260
Resolution (Å)	30-1.9 (2.0-1.9)
Mean redundancy	3.7 (3.0)
Unique reflections	37155
Completeness (%)	92.8 (89.5)
I/σ	16.4 (2.9)
R_{sym}	5.5 (34.0)
C. Refinement statistics	
R_F (%)	17.6
R_{free} (%)	20.1
rmsd from ideal	
Bond lengths (Å)	0.024
Bond angles (deg.)	1.126
Average B-factors (Å^2)	43.1
Ramachandran plot	
Most favored (%)	96.4
Allowed (%)	3.6
Generously allowed (%)	0
Disallowed (%)	0
D. Model content	
Monomers/ASU	1
Protein residues	2-195 211-296
Ligand	-
Water molecules	340

Table S1 Data collection and refinement statistics for the crystal structure of the EctD protein from *V. salexigens* in its iron-free form.

IV. Publication 4

The fourth publication is entitled ‘Crystal structure of the ectoine hydroxylase, a snapshot of the active site’. It was published in the ‘Journal of Biological Chemistry’ (doi: 10.1074/jbc.M114.576769).

Hoeppner, A., Widderich, N., Lenders, M., Bremer, E., and Smits, S.H. Crystal structure of the ectoine hydroxylase, a snapshot of the active site. *J Biol Chem*. (2014). 289: 29570-29583.

This paper reports the resolution of the crystal structure of the ectoine hydroxylase from the cold-adapted bacterium *Sphingopyxis alaskensis* in its apo-form, in complex with the iron(II) co-factor and in complex with the iron catalyst, the co-substrate 2-oxoglutarate and the reaction product 5-hydroxyectoine. Moreover, site-directed mutagenesis experiments demonstrate the role of different amino acid residues in ligand-binding. It thus reveals the architecture of the active site of EctD and increases the understanding of the structural-functional relationship of the ectoine hydroxylase family.

Personal contribution:
I performed the data base searches and the site-directed mutagenesis experiments, conducted overexpression and purification of EctD and its mutant derivatives and determined their kinetic properties. I further analyzed the experimental data (together with the other authors). Figures and Tables were prepared by me (and Sander Smits). I also contributed to the writing of the manuscript together with my PhD supervisor Erhard Bremer and with Sander Smits.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 289, NO. 43, pp. 29570–29583, October 24, 2014

Crystal Structure of the Ectoine Hydroxylase, a Snapshot of the Active Site*

Received for publication, April 25, 2014, and in revised form, August 27, 2014 Published, JBC Papers in Press, August 29, 2014, DOI 10.1074/jbc.M114.576769

Astrid Höppner[‡1], **Nils Widderich**[§¶1,2], **Michael Lenders**[‖], **Erhard Bremer**[§**3], **and Sander H. J. Smits**[‖4]

From the [‡]*X-ray Facility and Crystal Farm, Heinrich-Heine-University at Düsseldorf, D-40225 Düsseldorf, the* [§]*Department of Biology, Laboratory for Microbiology, Philipps-University at Marburg, D-35043 Marburg, the* [¶]*Max Planck Institute for Terrestrial Microbiology, Emeritus Group R. K. Thauer, D-35043 Marburg, the* [‖]*Institute of Biochemistry, Heinrich-Heine-University at Düsseldorf, D-40225 Düsseldorf, and the* [**]*LOEWE-Center for Synthetic Microbiology, Philipps-University at Marburg, D-35043 Marburg, Germany*

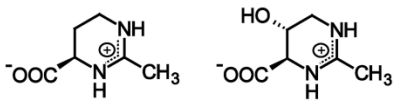
Background: 5-Hydroxyectoine is a compatible solute synthesized by the ectoine hydroxylase (EctD), a non-heme containing iron(II) and 2-oxoglutarate-dependent dioxygenase.
Results: The structure of EctD was solved in different forms.
Conclusion: The architecture of the catalytic core of EctD was revealed.
Significance: The crystal structure increases our understanding of the structural-functional relationship in an evolutionary conserved group of enzymes.

Ectoine and its derivative 5-hydroxyectoine are compatible solutes that are widely synthesized by bacteria to cope physiologically with osmotic stress. They also serve as chemical chaperones and maintain the functionality of macromolecules. 5-Hydroxyectoine is produced from ectoine through a stereo-specific hydroxylation, an enzymatic reaction catalyzed by the ectoine hydroxylase (EctD). The EctD protein is a member of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily and is evolutionarily well conserved. We studied the ectoine hydroxylase from the cold-adapted marine ultra-microbacterium *Sphingopyxis alaskensis* (*Sa*) and found that the purified *Sa*EctD protein is a homodimer in solution. We determined the *Sa*EctD crystal structure in its apo-form, complexed with the iron catalyst, and in a form that contained iron, the co-substrate 2-oxoglutarate, and the reaction product of EctD, 5-hydroxyectoine. The iron and 2-oxoglutarate ligands are bound within the EctD active site in a fashion similar to that found in other members of the dioxygenase superfamily. 5-Hydroxyectoine, however, is coordinated by EctD in manner different from that found in high affinity solute receptor proteins operating in conjunction with microbial import systems for ectoines. Our crystallographic analysis provides a detailed view into the active site of the ectoine hydroxylase and exposes an intricate network of interactions between the enzyme and its ligands that collectively ensure the hydroxylation of the ectoine substrate in a position- and stereo-specific manner.

Cells of all three kingdoms of life exploit compatible solutes as protectants against water loss that ensues upon their exposure to high osmolarity surroundings (1, 2). These organic compounds can be accumulated to exceedingly high intracellular concentrations without exerting negative side effects on the biochemistry of the cell. They thus serve as physiologically compliant water-attracting osmolytes to maintain an adequate level of cellular hydration (3). The physicochemical attributes of compatible solutes that make their physiological function(s) possible largely stem from their negative interactions with the protein backbone that in turn leads to their preferential exclusion from protein surfaces (4). As a consequence of the resulting imbalance of the distribution of compatible solutes in the cell water, thermodynamics favor the native conformation of proteins and the functionality of other macromolecules under otherwise function-disrupting conditions (5, 6). The term "chemical chaperone" has been coined in the literature to reflect these beneficial traits (7).

Ectoine ((4*S*)-2-methyl-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid) (Fig. 1) (8) is a well recognized compatible solute and microbial stress protectant (9, 10). Its function-preserving (10–15) and anti-inflammatory properties fostered substantial interest for its industrial scale production (9, 16) and exploitation in various biotechnological applications, in skin-care products (9, 10, 16–18), and in potential medical uses (19–21). Ectoine is synthesized widely by members of bacteria (10, 16, 22) as an adaptive response to osmotic stress (16, 23–29) and growth-restricting extremes in high and low temperatures (25, 30). Synthesis of ectoine relies on the supply of the precursor L-aspartate-β-semialdehyde and is carried out through the enzymatic

* This work was supported in part by Deutsche Forschungsgemeinschaft Grant SFB 987, the LOEWE program of the State of Hessen (via the Center for Synthetic Microbiology, Marburg), the Max-Planck Institute for Terrestrial Microbiology (Marburg) through the Emeritus Group of R. K. Thauer, a contribution by the Fonds der Chemischen Industrie, by the Heinrich-Heine-University Düsseldorf and its Institute of Biochemistry, and by the initiative "Fit for Excellence" of the Heinrich-Heine-University.
The atomic coordinates and structure factors (codes 4Q5O, 4MHR, and 4MHU) have been deposited in the Protein Data Bank (http://wwpdb.org/).
[1] Both authors contributed equally to this work.
[2] Recipient of a Ph.D. fellowship from the International Max Planck Research School for Environmental, Cellular, and Molecular Microbiology, Marburg, Germany.
[3] To whom correspondence may be addressed: Philipps-University Marburg, Karl-von-Frisch Str. 8, D-35043 Marburg, Germany. Tel.: 49-6421-282-1529; Fax: 49-6421-282-8979; E-mail: bremer@staff.uni-marburg.de.
[4] To whom correspondence may be addressed: Heinrich-Heine-University, Institute of Biochemistry, Universitätsstr. 1, D-40225 Düsseldorf, Germany. Tel.: 49-2118112647; Fax: 49-211-8115310; E-mail: Sander.Smits@hhu.de.



ectoine 5-hydroxyectoine

FIGURE 1. **Chemical structures of ectoine and 5-hydroxyectoine.**

activities of L-2,4-diaminobutyrate transaminase (EctB; EC 2.6.1.76), 2,4-diaminobutyrate acetyltransferase (EctA; EC 2.3.1.178), and ectoine synthase (EctC; EC 4.2.1.108) (22, 31, 32).

A substantial group of ectoine producers (22) also synthesize a derivative in which a hydroxy group is attached to the pro-*S* hydrogen at position C5 in the tetrahydropyrimidine ring of ectoine in a stereo-specific manner, thereby yielding 5-hydroxyectoine ((4*S*,5*S*)-5-hydroxy-2-methyl-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid (Fig. 1) (33), a reaction carried out by the ectoine hydroxylase (EctD; EC 1.14.11) (27, 30, 34). Despite their closely related chemical structures, 5-hydroxyectoine is often superior to its precursor molecule ectoine in protecting microorganisms against environmentally imposed stresses (26, 30) and in preserving the functionality of macromolecules and cells (11–15, 35, 36). This makes a better understanding of the ectoine hydroxylase desirable, both in terms of basic science and for its practical use in biotechnological processes.

The ectoine hydroxylase is a member of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily, a group of versatile biocatalysts that not only carry out hydroxylation reactions but also dimethylations, desaturations, cyclizations, ring expansions, epimerizations, and halogenations (37–41). Although the type of substrates used by these dioxygenases varies considerably, common enzyme reaction mechanisms are observed. Most members of this superfamily couple the two-electron oxidation of their substrates with the reaction of oxygen and 2-oxoglutarate (38, 39, 42). The ectoine hydroxylase adheres to this general reaction scheme (27, 43).

Despite their varied enzyme activities, non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases are structurally closely related and typically consist of a double-stranded β-helix (DSBH)[5] core, the so-called jelly roll or cupin fold (37–40). Crystallographic analysis of the ectoine hydroxylase from the salt-tolerant moderate halophile *Virgibacillus salexigens* (*Vs*) in its iron-bound (44) and iron-free (22) forms revealed that the DSBH in the *Vs*EctD protein is formed by four-stranded antiparallel β-sheets arranged in form of a β-sandwich; it is decorated with and stabilized by a number of α-helices (22, 44). However, both available crystal structures of the *Vs*EctD protein lack the co-substrate 2-oxoglutarate and the substrate ectoine, and consequently, our understanding of the structure and architecture of the active site of this enzyme is incomplete.

[5] The abbreviations used are: DSBH, double-stranded β-helix; r.m.s.d., root mean square deviation; ASU, asymmetric unit; TES, 2-{[2-hydroxy-1,1-bis(hydroxymethyl)ethyl]amino}ethanesulfonic acid; PDB, Protein Data Bank; MALS, multiangle light scattering.

Recently, we have biochemically characterized six ectoine hydroxylases from microorganisms that are able to colonize habitats with extremes in salinity, growth temperature, and pH values (22). These six enzymes possess kinetic parameters similar to those of the previously studied ectoine hydroxylases from *V. salexigens* and *Streptomyces coelicolor* (26, 27), but differences in their ability to withstand the inhibiting effects of high salt concentrations and high temperature were noted for some of them (22). These enzymes might thus be better suited than the *V. salexigens* EctD protein to obtain a crystal structure containing all ligands.

Here, we report the crystal structures of the ectoine hydroxylase from the cold-adapted marine ultra-microbacterium *Sphingopyxis alaskensis* (45), *Sa*EctD (22, 46), in its apo-form, in complex with iron and in a form that contains the iron catalyst, the co-substrate 2-oxoglutarate and the product of the EctD-catalyzed enzyme reaction, 5-hydroxyectoine. Combined with previous molecular dynamics simulations (44), the crystallographic analysis presented here and our site-directed mutagenesis experiments revealed the architecture of the active site of this evolutionarily well conserved group of enzyme.

EXPERIMENTAL PROCEDURES

Chemicals—Ectoine and 5-hydroxyectoine were kindly provided by Dr. Thomas Lentzen and Dr. Irina Bagyan (bitop AG, Witten, Germany). 2-Oxoglutarate (disodium salt) was purchased from Sigma. Anhydrotetracycline hydrochloride, desthiobiotin, and Strep-Tactin Superflow chromatography material were obtained from IBA GmbH (Göttingen, Germany).

Bacterial Strains, Media, and Growth—The *Escherichia coli* strain DH5α (Invitrogen) was used for the propagation of the *ectD*$^+$ plasmid pMP40 (22, 46). It was grown at 37 °C in Luria-Bertani (LB) liquid medium or on LB agar plates containing ampicillin (100 μg ml^{-1}) to select for the presence of the plasmid. Heterologous overproduction of the *S. alaskensis* EctD protein was carried out in the *E. coli* B strain BL21 (pMP40) in minimal medium A (47) containing 0.5% (w/v) glucose as carbon source, 0.5% (w/v) casamino acids, 1 mM $MgSO_4$, and 3 mM thiamine.

Site-directed Mutagenesis of the Sa-ectD Gene—To assess the contribution of amino acids predicted by structural analysis to be involved in the binding of the 5-hydroxyectoine molecule, or predicted to contribute to interactions between the *Sa*EctD monomers, we performed site-directed mutagenesis. For these experiments, we used the QuikChange Lightning mutagenesis kit (Stratagene, La Jolla, CA) and custom-synthesized mutagenic DNA primers; the DNA primers had a typical length of 25 nucleotides and were purchased from Biomers (Ulm, Germany). To ensure the presence of the desired mutation(s) and the absence of unwanted alterations within the mutagenized *ectD* gene, the DNA sequence of the entire coding region was determined (Eurofins MWG GmbH, Ebersberg, Germany) for each *ectD* mutant allele. DNA of plasmid pMP40 carrying the *Sa-ectD* wild-type gene was used as the template for the site-directed mutagenesis experiments. The following *Sa-ectD* gene variants were isolated: pWN20 (Gln-127 to Ala (CAG to GCG)), pWN21 (Arg-139 to Ala (AGG to GCG)), pWN22 (Glu-140 to Ala (GAG to GCG)), pWN23 (Thr-149 to Ala (ACC to

GCC)), pWN24 (Trp-150 to Ala (TGG to GCG)), pWN25 (Arg-280 to Ala (CGC to GCC)), and the double mutant pWN26 (Arg-139 to Ala (AGG to GCG) and Glu-140 to Ala (GAG to GCG)). These *Sa*EctD variants were overproduced in the *E. coli* B strain BL21 and purified by affinity chromatography. They were functionally assessed by the previously described ectoine hydroxylase enzyme assay (22, 46).

Overproduction and Purification of the Recombinant EctD Protein—Plasmid pMP40 is a derivative of the expression vector pASK-IBA3 (IBA, Göttingen, Germany) in which the transcription of the *ectD* gene (accession number YP_617990) from *S. alaskensis* (45) is positioned under the control of the TetR-responsive and anhydrotetracycline hydrochloride-inducible *tet* promoter carried by the backbone of the pASK-IBA3 expression vector (43, 46). In this plasmid, a nine-amino acid *Strep*-tag-II affinity peptide encoded by a short synthetic DNA sequence is fused to the 3′ end of the *ectD*-coding region to allow the purification of the encoded *Sa*EctD-*Strep*-tag-II protein by affinity chromatography using Strep-Tactin Superflow material. Overproduction and purification of this recombinant protein for crystallization experiments was carried out as described (46). 200–300 mg of *Sa*EctD-*Strep*-tag-II protein/liter of culture was routinely obtained, and the protein was concentrated with Vivaspin 6 columns (Sartorius Stedim Biotech GmbH, Göttingen, Germany) to about 10 mg ml^{-1} (in 20 mM TES buffer, pH 7.5, 80 mM NaCl) before it was used for crystallization trials. The purity of the obtained *Sa*EctD-*Strep*-tag-II protein preparations was assessed on 12% SDS-polyacrylamide gels, and their iron content was determined with the aid of a colorimetric assay (27, 48).

Ectoine Hydroxylase Enzyme Activity Assay—Ectoine hydroxylase enzyme activity of the *Sa*EctD-*Strep*-tag-II protein was determined under previously described buffer and assay conditions (26, 27, 46). Specifically, we monitored the EctD-mediated conversion of ectoine into 5-hydroxyectoine in a 30-μl reaction volume that contained 6 mM ectoine, 10 mM 2-oxoglutarate, 1 mM $FeSO_4$, 20 mM TES, pH 8.0, and 100 mM KCl. 15 μg of either the EctD wild-type protein or the particular *Sa*EctD mutant protein under study was added to the reaction mixture. Each assay was run for 10 min at 40 °C under vigorous aeration. The enzyme reaction was stopped by adding 30 μl of acetonitrile (100%) to the reaction vessel. The substrate ectoine and the reaction product 5-hydroxyectoine were analytically separated from each other on a GROM-SIL Amino-1PR column (125 × 4 mm with a particle size of 3 μm); it was purchased from GROM (Rottenburg-Hailfingen, Germany). The conversion of ectoine to 5-hydroxyectoine by the purified *Sa*EctD-*Strep*-tag-II protein was monitored with a UV-visible detector integrated into a FPLC system (SYKAM, Fürstenfeldbruck, Germany) by measuring the absorbance of these solutes at a wavelength of 210 nm (22, 24, 27). The amount of ectoine remaining in the enzyme assay and the amount of 5-hydroxyectoine formed by the wild-type and mutant EctD proteins were quantified using the ChromStar 7.0 software package (SYKAM, Fürstenfeldbruck, Germany). The data shown for each mutant were derived from two independent *Sa*EctD preparations, and each *Sa*EctD protein solution was assayed three times for its enzyme activity.

Determination of the Oligomeric State of Purified SaEctD and Mutant Derivatives by Conventional Size-exclusion Chromatography and HPLC-MALS—To determine the oligomeric state of the *Sa*EctD-*Strep*-tag-II protein in solution, we used conventional size-exclusion chromatography and high performance liquid chromatography coupled to multiangle light scattering detection (HPLC-MALS). Size-exclusion chromatography was performed with a HiLoad 16/600 Superdex 200pg column (GE Healthcare) equilibrated and run in a 20 mM TES buffer, pH 7.5, containing 150 mM NaCl under conditions described previously (22). The size-exclusion column was standardized with a gel filtration markers kit (Sigma). For HPLC-MALS analysis, a Bio SEC-5 HPLC column (Agilent Technologies Deutschland GmbH, Böblingen, Germany) with a pore size of 300 Å was equilibrated with 20 mM TES, pH 7.5, 80 mM NaCl for HPLC using a system from Agilent Technologies connected to a triple-angle light-scattering detector (miniDAWN TREOS, Wyatt Technology Europe GmbH, Dernbach, Germany) followed by a differential refractive index detector (OPTILab T-rEX, Wyatt Technology). Typically, 100 μl of purified *Sa*EctD (2.0 mg ml^{-1}) was loaded onto the Bio SEC-5 HPLC column, and the obtained data were analyzed with the ASTRA software package (Wyatt Technology).

Crystallization of the SaEctD Protein—For apo-*Sa*EctD, the crystallization trials were performed using the sitting-drop vapor diffusion method at 20 °C. 1.5 μl of the homogeneous protein solution of *Sa*EctD-*Strep*-tag-II (10 mg ml^{-1} in 20 mM TES, pH 7.5, 80 mM NaCl) was mixed with 1.5 μl reservoir solution containing 100 mM MES, pH 6.0, 200 mM calcium acetate, 30% (w/v) PEG 400, and 1.5 mM *n*-dodecyl-*N*,*N*-dimethylglycine and equilibrated over th 300 μl reservoir solution. Crystals grew within 6–12 days to their final size of around 30 × 30 × 50 μm^3. For Fe-*Sa*EctD-*Strep*-tag-II in complex with Fe^{2+}, the *Sa*EctD-*Strep*-tag-II enzyme in complex with its iron catalyst was crystallized as described above for apo-*Sa*EctD-*Strep*-tag-II protein except that the protein solution was premixed with 100 mM $Fe(II)Cl_2$ to a final concentration of 4 mM and incubated on ice for 10–15 min. In addition, the solution contained 3.5 mM *n*-dodecyl-*N*,*N*-dimethylglycine. Crystals grew within 6–12 days at 20 °C to their final size of around 40 × 40 × 180 μm^3. For Fe-*Sa*EctD-*Strep*-tag-II in complex with 2-oxoglutarate and 5-hydroxyectoine, the *Sa*EctD-*Strep*-tag-II protein was mixed with $Fe(II)Cl_2$ as described above, and 2-oxoglutarate was subsequently added to the protein solution to a final concentration of 40 mM. After 30 min of incubation on ice, 5-hydroxyectoine was added to the solution to a final concentration of 40 mM. This mixture was then incubated for 1 h on ice before crystallization trials were conducted. *Sa*EctD crystals were grown by mixing 1.5 μl of protein solution with 1.5 μl of reservoir containing 100 mM MES, pH 6.0, 200 mM calcium acetate, 30% (w/v) PEG 400, and 25 mM *n*-octyl-β-D-glycoside; they grew within 6–12 days to their final size of around 40 × 40 × 70 μm^3. All crystals were cryoprotected by carefully adding 1 μl of 100% glycerol to the crystallization drop before the crystals were frozen in liquid nitrogen.

Data Processing and Structure Determination—Data sets were collected from a single crystal obtained from the various crystallization trials (apo-*Sa*EctD, Fe-*Sa*EctD, and Fe-*Sa*EctD/

2-oxoglutarate/5-hydroxyectoine) at the ERSF beamline ID23eh2 (Grenoble, France) at 100 K. These data sets were processed using the XDS package (49) and scaled with XSCALE (50). Initial phases were obtained by molecular replacement using the program PHASER (51) with the crystal structure of the *V. salexigens* EctD protein (PDB code 3EMR) as the template (44). Model building and refinement were performed using COOT (52) and REFMAC5 (53). Data refinement statistics and model content are summarized in Table 1. The structures were deposited at the Protein Data Bank (Brookhaven, NY) under the following accession codes: 4MHR for the apo-*Sa*EctD protein, 4MHU for the Fe-*Sa*EctD protein, and 4Q5O for Fe-*Sa*EctD/2-oxoglutarate/5-hydroxyectoine crystal structure.

Database Searches for EctD-related Proteins—We used the amino acid sequence (accession number YP_617990) from *S. alaskensis* (45) to search microbial genomes for EctD-related proteins in the databases of the National Center for Biotechnology Information (NCBI) (www.ncbi.nlm.nih.gov) and of the United States Department of Energy Joint Genomic Institute) (jgi.doe.gov). Amino acid sequences of EctD-type proteins were aligned as described previously (22, 43).

Figure Preparation of Crystal Structures—Figures of the crystal structure of the *Sa*EctD protein were prepared using the PyMOL software package and Chimera.

RESULTS

Biochemical Properties of the Ectoine Hydroxylase from S. alaskensis and its Oligomeric State in Solution—The ectoine hydroxylase studied here originates from *S. alaskensis*, a microorganism that is well adapted to a life in chilly ocean water systems (45). Despite this permanently cold habitat of the producer microorganism, the *Sa*EctD enzyme itself is thermostable (22). The kinetic properties of the *Sa*EctD protein (with K_m values of 9.8 ± 0.5 and 2.7 ± 0.3 mM for the substrate ectoine and the co-substrate 2-oxoglutarate, respectively, and a catalytic efficiency (k_{cat}/K_m) of 0.12 $mM^{-1}\ s^{-1}$) resemble those of other biochemically studied ectoine hydroxylases (22, 26, 27).

We used for our studies a recombinant *Sa*EctD-*Strep*-tag-II protein that carries a nine-amino acid tag at its carboxy terminus to allow affinity purification on Strep-Tactin Superflow material (22, 46). To determine the oligomeric state and to validate the monodispersity of the EctD protein in solution, we carried out a HPLC-MALS analysis. The normalized elution profiles from the UV, refractive index, and light scattering detectors revealed one symmetric peak indicating that the protein solution was homogeneous and monodisperse. After determining protein concentration by the refractive index, a molecular mass of 70.73 ± 1.1 kDa was obtained for the *Sa*EctD-*Strep*-tag-II protein (Fig. 2). This molecular mass corresponds very well to the theoretical molecular mass of a dimer of the recombinant protein (calculated molecular mass of the monomer, including the *Strep*-tag-II affinity peptide: 35.29 kDa) (46). We therefore conclude that the ectoine hydroxylase from *S. alaskensis* is a dimer in solution, a conclusion that is in agreement with the suggested quaternary structures of six other EctD proteins that have been assessed by conventional size exclusion chromatography (22).

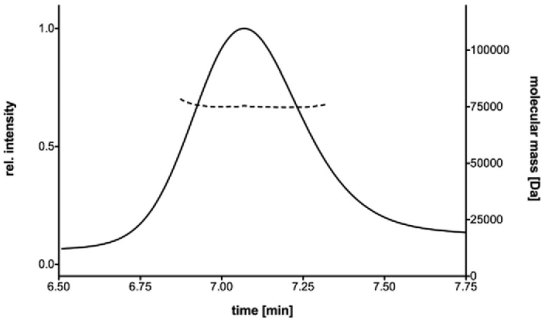


FIGURE 2. **Determination of the oligomeric state of the purified *Sa*EctD-*Strep*-tag-II protein by HPLC-MALS analysis.** *Black line,* normalized refractive index detector signal; *black dotted line,* calculated protein mass.

Crystallization of the SaEctD-Strep-tag-II Protein—By slightly varying the previously used overproduction conditions for the *Sa*EctD-*Strep*-tag-II protein in *E. coli* (22), we were able to increase the amounts of the recovered *Sa*EctD-*Strep*-tag-II protein/liter of culture about 10-fold to 200–300 mg (46). However, although the modified overproduction procedure aided the provision of large amounts of protein for crystallization trials, this also reduced the iron content (44) of the *Sa*EctD-*Strep*-tag-II protein preparation to about 0.13 mol of iron/mol of protein and thereby caused a strong decrease in ectoine hydroxylase activity. By adding $FeSO_4$ or $FeCl_2$ to the purified protein solution, the ectoine hydroxylase activity could be restored, and this allowed us to obtain crystals of the *Sa*EctD-*Strep*-tag-II protein in its apo- and iron-bound forms. The crystals of the apo-*Sa*EctD and Fe-*Sa*EctD proteins diffracted to a resolution of 2.1 and 2.64 Å, respectively.

Co-crystallization experiments of *Sa*EctD with its substrate ectoine were not successful, despite that a large number of well diffracting crystals were obtained; however, none of these contained the ectoine ligand as revealed by x-ray analysis. The co-crystallization of a protein with its enzymatic reaction product has been successfully used to obtain informative crystal structures (54). When applied to the *Sa*EctD protein, it yielded crystals with good diffraction properties, but none of them showed any electron density for 5-hydroxyectoine. We then turned to an approach in which we added iron, 2-oxoglutarate, and 5-hydroxyectoine in a stepwise manner to the *Sa*EctD-*Strep*-tag-II protein in which we incubated the protein solution containing a high concentration of the ligands for at least 1.5 h on ice before it was used for crystallization trials. In this way, we obtained crystals that diffracted to a resolution of 2.56 Å and that contained the iron ligand, the co-substrate 2-oxoglutarate, and the EctD-enzyme reaction product 5-hydroxyectoine.

All three crystal structures of the *Sa*EctD protein were solved by molecular replacement using the crystal structure of the *V. salexigens* (*Vs*EctD) ectoine hydroxylase (PDB code 3EMR) (44) as the search model. A summary of the data collection statistics, refinement details, and model content is given in Table 1.

Overall Fold of the SaEctD Protein—The amino acid sequences of the *Vs*EctD and *Sa*EctD possess an amino acid

TABLE 1
Data collection and refinement statistics for the apo- and iron-bound forms of the ectoine hydroxylase and of the iron-, 2-oxoglutarate-, and 5-hydroxyectoine-bound forms of the *Sa*EctD protein

	Apo-*Sa*EctD	Fe-*Sa*EctD	Fe-*Sa*EctD/2-oxoglutarate/5-hydroxyectoine
Crystal parameters at 100 K			
Space group	C222	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters			
a, *b*, *c* (Å)	83.48, 86.51, 95.34	78.16, 87.52, 96.05	81.00, 87.08, 94.88
$\alpha = \beta = \gamma$ (°)	90	90	90
Data collection and processing			
Wavelength (Å)	0.87260	0.87260	0.87260
Resolution (Å)	30-2.1 (2.2-2.1)	30-2.64 (2.8-2.64)	30-2.56 (2.8–2.56)
Mean redundancy	4.2 (4.1)	4.9 (5.0)	3.9 (3.8)
Unique reflections	20,251	18,652	18,360
Completeness (%)	99.7 (99.8)	99.6 (99.9)	99.8 (99.8)
I/σ	15.1 (2.8)	19.8 (2.9)	14.8 (1.8)
R_{sym} (%)	6.2 (49.9)	5.8 (58.9)	7.3 (56.2)
Refinement statistics			
R_F (%)	16.7	20.6	20.6
R_{free} (%)	21.8	26.7	27.8
r.m.s.d. from ideal			
Bond lengths (Å)	0.019	0.0115	0.009
Bond angles (°)	1.8	1.5	1.3
Average B-factors (Å^2)	43.1	63.8	54.4
Ramachandran plot			
Most favored (%)	96.3	93.4	92.1
Allowed (%)	3	5.1	6.6
Generously allowed (%)	0.7		
Disallowed (%)		1.5	1.4
Model content			
Monomers/ASU	1	2	2
Protein residues	2–192, 210–301	2–192, 210–301	2–192, 210–301
Ligand		2 iron, 1 D9G	2 iron, 2 AKG, 2 6CS
Water molecules	108		

sequence identity of 51% (for an alignment see Fig. 3). An overall comparison of the three newly determined *Sa*EctD crystal structures with that of the previously determined structure of *Vs*EctD revealed a high degree of identity with an r.m.s.d. that ranges from 1.3 to 1.6 Å over 279 C-α atoms. An even lower r.m.s.d. of 0.5–0.8 Å was found when the three different *Sa*EctD structures were compared with each other. Hence, neither the binding of the catalytically critical iron ligand nor of the substrate 2-oxoglutarate and 5-hydroxyectoine triggered a strong change in the overall conformation of the EctD protein.

Because the monomers of all three *Sa*EctD structures are nearly identical in overall shape, we describe in the following section only the overall fold for the apo-*Sa*EctD protein. The *Sa*EctD crystal structure consists of a double-stranded β-helix core surrounded and stabilized by a number of α-helices (Fig. 4*A*). This core, also known as the jelly roll or cupin fold (38, 39), is formed by two four-stranded anti-parallel β-sheets that are arranged in the form of a β-sandwich (Fig. 4*A*). This type of architecture has previously been observed not only for the *Vs*EctD protein (22, 44) but also for many other non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases (38–40). Examples are the EctD-related structures of the human phytanoyl-CoA hydroxylases PhyH (PDB code 2A1X; r.m.s.d. of 2.6 Å over 252 C-α atoms when compared with the *Sa*EctD structure) (55) and PhyHD1A (PDB code 2OPW r.m.s.d. of 2.2 Å over 286 C-α atoms) (56) and the halogenases SyrB2 from *Pseudomonas syringae* (PDB code 2FCV; r.m.s.d. of 2.9 Å over 299 C-α atoms) (57), CytC3 from a *Streptomyces* soil isolate (PDB code 3GJA; r.m.s.d. of 3.0 Å over 204 C-α atoms) (58), and CurA from the cyanobacterium *Lyngbya majuscula* (PDB code 3NNM; r.m.s.d. of 3.3 Å over 222 C-α atoms) (59). This structural comparison shows that the core of the EctD protein, its cupin-type fold, is very similar to these proteins; however, the rest of EctD differs from the structurally compared proteins and thereby yields higher r.m.s.d. values.

In each of the three *Sa*EctD crystal structures, a large loop (amino acids 191–210) is disordered and hence not visible in the electron density map; the corresponding region was therefore not included in the *Sa*EctD models. By amino acid sequence alignment (Fig. 3), this loop corresponds to residues 195–211 in the *Vs*EctD ectoine hydroxylase, and it is also disordered in both crystal structures of this protein (22, 44). Notably, in the ligand-free crystal structures of the phytanoyl-CoA hydroxylases PhyH (55), the asparagine hydroxylase AsnO from *S. coelicolor* (60), the L-arginine oxygenase VioC from *Streptomyces vinaceus* (61), and the taurine dioxygenases from *E. coli* and *Pseudomonas putida* (62), a similar loop is disordered as well and became only visible in crystal structures with bound substrates. It is thought that this mobile loop functions as a lid that shields the enzyme reaction chamber during catalysis from the solvent.

Dimer Interface of the SaEctD Protein—The three crystal structures of the ectoine hydroxylase from *S. alaskensis* displayed a different crystal packing (as observed by the different symmetry of the crystals) and protein composition in the asymmetric unit (ASU) of the crystals. In apo-*Sa*EctD, the asymmetric unit contains a monomer, whereas the ASUs of Fe-*Sa*EctD and Fe-*Sa*EctD/2-oxoglutarate/5-hydroxyectoine both include dimers. The monomers in the latter dimer are however differently oriented (Fig. 5). Because the *Sa*EctD protein is a dimer in solution (see above and Fig. 2), we inspected the crystal packing and analyzed the monomer/monomer interactions to elucidate

FIGURE 3. **Alignment of the amino acid sequences of 10 EctD-type proteins.** The *S. alaskensis* ectoine hydroxylase was used as the search query, and the EctD-type proteins from *Sphingobium japonicum* (WP_006964700) and *Sphingobium baderi* (WP_021243884) were found as the most closely related proteins to the *Sa*EctD (YP_617990) enzyme. The EctD enzymes from *Acidiphilium cryptum* (AER00258), *Alkalilimnicola ehrlichii* (AER00257), *Paenibacillus lautus* (ACX67869), *Virgibacillus salexigens* (AAY29689), *Halomonas elongata* (YP_003899077), *S. coelicolor* (Q93RV9), and *Pseudomonas stutzeri* (WP_011911424) have been functionally characterized and are *bona fide* ectoine hydroxylases. Residues involved in the binding of the iron catalyst are marked in *red*; those that mediate the binding of the 2-oxoglutarate co-substrate are labeled in *green*, and residues contacting the 5-hydroxyectoine molecule are marked in *blue*. Residues His-144 and Asp-146 of the *Sa*EctD protein are involved both in the binding of the iron catalyst and also of the 5-hydroxyectoine molecule. The string of the 17 amino acids that serves as the consensus sequence of ectoine hydroxylases is labeled in *yellow*.

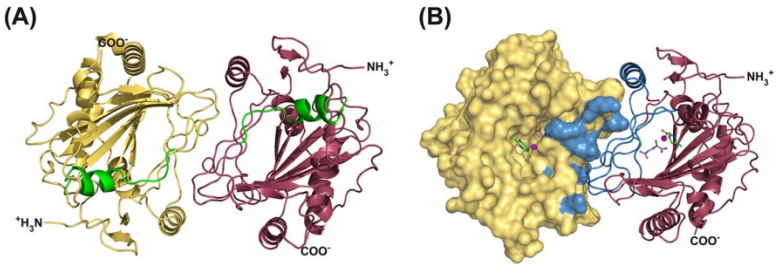


FIGURE 4. **Dimer structure of the *Sa*EctD-*Strep*-tag-II protein in complex with iron, 2-oxoglutarate, and 5-hydroxyectoine.** *A, ribbon* representation of the crystal structure of the *Sa*EctD-*Strep*-tag-II protein in complex with iron, 2-oxoglutarate, and 5-hydroxyectoine (PDB code 4Q5O); the ligands have been omitted for simplicity. The two monomers are differently colored, and the region corresponding to the EctD signature sequence motif is highlighted in *green,* and the N and C termini of the protein are indicated. *B,* representation of the monomer/monomer interface of the *Sa*EctD-*Strep*-tag-II dimer in complex with all three substrates. One of the monomers is shown in surface representation, and the second monomer is represented in a *ribbon* format. Residues and regions that interact in the dimer are highlighted in *blue*. The iron catalyst is represented as a *sphere* (in *magenta*) and the co-substrate 2-oxoglutarate and the EctD enzyme reaction product, 5-hydroxyectoine, are shown in *ball and stick* representation.

which dimer in the crystals might represent the physiologically relevant dimer (Fig. 5). The monomers in the ASUs of apo-*Sa*EctD form also a dimer assembly through crystal symmetry, similar to that found in the *Sa*EctD/2-oxoglutarate/5-hydroxyectoine crystal structure. The monomers in the Fe-*Sa*EctD dimer are slightly shifted toward each other when

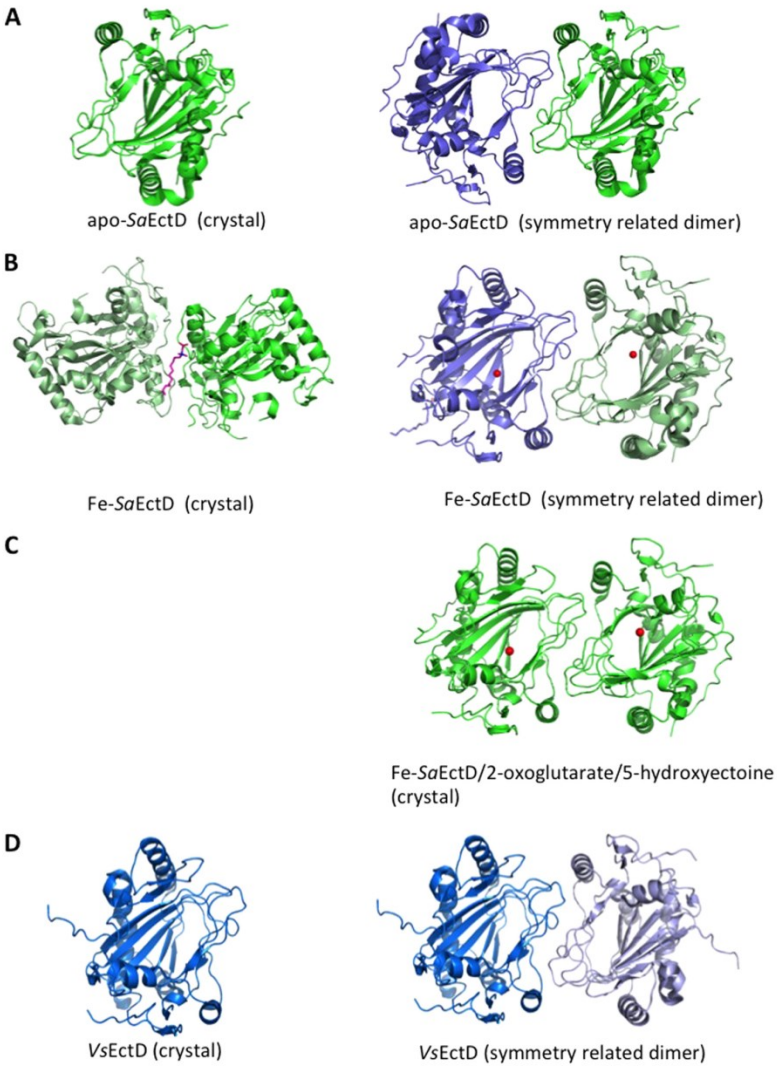


FIGURE 5. **Dimer assemblies of the *Sa*EctD protein.** Shown are the compositions of the ASUs of the apo-*Sa*EctD, Fe-*Sa*EctD, and Fe-*Sa*EctD/2-oxoglutarate/5-hydroxyectoine crystals, and the associated dimer of the *Sa*EctD protein obtained by symmetry-related molecules. *A,* apo-*Sa*EctD was crystallized containing a monomer in the ASU (highlighted in *green*). The corresponding dimer was obtained by highlighting the symmetry-related molecules. The second monomer of the dimer is shown in *blue*. *B,* Fe-*Sa*EctD contained a dimer in the ASU (the single monomers are shown in *light* and *dark green*) and an extra *n*-dodecyl-*N*,*N*-dimethylglycine (highlighted in *red*) captured from the crystallization solution; it was bound between the two monomers. The dimer found by symmetry-related molecules is highlighted in *green* and *blue* in the *right panel*. The iron molecules are shown as *red spheres*. *C,* Fe-*Sa*EctD/2-oxoglutarate/5-hydroxyectoine crystals contained a dimer in the ASU similar to the other symmetry related dimers. The iron molecules were highlighted as *red spheres*. For clarity we omitted the 2-oxoglutarate and 5-hydroxyectoine from the figure. *D,* crystal structure of the *Sa*EctD-related *V. salexigens* EctD protein (PDB code 3EMR) (44) revealed a monomer in the ASU shown in *blue*. The iron molecules are highlighted in *red*. The symmetry related dimer is shown in the *right panel*.

one compares them to the *Sa*EctD/2-oxoglutarate/5-hydroxyectoine dimer. However, this difference in the Fe-*Sa*EctD structure is likely due to the extra *n*-dodecyl-*N*,*N*-dimethylglycine molecule that was present in the crystallization solution and that is bound in between the two monomers. This leads to a different crystal packing in the ASU. In contrast, when looking at the Fe-*Sa*EctD dimer composed of the symmetry-related monomers (Fig. 5*B*, *right panel*), we found the same arrangement in the ASU of the *Sa*EctD/2-oxoglutarate/5-hydroxyectoine crystal (Fig. 5*C*). The only other member of the EctD superfamily that has been so far crystallized is the *Vs*EctD protein in forms that either possess or lack the iron catalyst (22, 44). In these structures, the ASU contains a monomer, although gel filtration analysis showed that the *Vs*EctD protein is a dimer in solution (22). We re-analyzed the crystal packing of the *Vs*EctD protein and found a similar dimer as observed in the *Sa*EctD/

2-oxoglutarate/5-hydroxyectoine crystal structure presented here (Fig. 5*D*).

To ensure that the dimer observed in the *Sa*EctD/2-oxoglutarate/5-hydroxyectoine crystal structure is also a dimer in solution, we studied its quaternary composition under the very same buffer and substrate conditions that were used in the co-crystallization experiments by analyzing these samples using size-exclusion chromatography-MALS. The molecular mass determined for the *Sa*EctD-*Strep*-tag-II protein by this experiment was 75.02 ± 1.8 kDa, which corresponds to a dimeric state of the protein (calculated theoretical mass of the iron, 2-oxoglutarate, and 5-hydroxyectoine bound dimer, 71.3 kDa). We then analyzed the dimer interface using the PDBePISA software (63). We found 10 potential monomer/monomer interaction regions (Table 2) and observed that the involved amino acids are all located in loop areas pointing from one monomer to the other monomer (Fig. 4*B*). However, inspection of the aforementioned loop regions implicated in forming the dimer interface revealed that their amino acid sequences are only moderately conserved, as judged by consulting a previous alignment of 433 EctD-type proteins (22, 43).

Residues Arg-139 and Glu-140 seemed to make most contributions to dimer formation (Table 2). To investigate their potential role for the quaternary structure and their possible corresponding influence on enzyme activity, we targeted these residues of *Sa*EctD by site-directed mutagenesis. We replaced them individually by Ala residues and also constructed the respective double mutant. Each of these three EctD variants was fully catalytically active (Table 3). Contrary to expectations, each of these EctD proteins still formed dimers in solution, as assessed by conventional size-exclusion chromatography This is documented for the double mutant EctD protein (Arg-139/Ala and Glu-140/Ala) in Fig. 6.

TABLE 2

Monomer/monomer interactions in the *Sa*EctD protein in complex with Fe-2-oxoglutarate and 5-hydroxyectoine

The amino acids involved in the dimer interface are listed, and the distance and type of interaction ((H), hydrogen bond; (S), salt bridge) are indicated.

Monomer A	Monomer B	Distance (type)
Glu-140 (OE1)	Arg-139 (NH_2)	2.6 Å (H)
Glu-140 (OE1)	Arg-139 (NE)	3.7 Å (S)
Glu-140 (OE1	Arg-139 (NH1)	3.7 Å (S)
Glu-140 (OE1)	Arg-139 (NH_2)	2.5 Å (S)
His-174 (NE2)	Ile-226 (O)	3.3 Å (H)
Glu-214 (OE2)	Lys-137 (NZ)	3.6 Å (S)
Gly-249 (O)	Asn-248 (ND2)	3.3 Å (H)
Ile-251 (O)	Trp-143 (NE1)	3.1 Å (H)

TABLE 3

Conversion of ectoine into 5-hydroxyectoine by *Sa*EctD mutants

The conversion of ectoine into 5-hydroxyectoine by the *Sa*EctD protein was monitored in a reaction that contained 6 mM ectoine as the substrate, 10 mM 2-oxoglutarate as the co-substrate, and 15 μg of the EctD protein under study. The amount of 5-hydroxyectoine was measured after 10 min.

EctD variant	5-Hydroxyectoine	Proposed role in
	mM	
Wild type	4.93 ± 0.26	
Q127A	0.04 ± 0.03	Ectoine binding
T149A	0.17 ± 0.05	Ectoine binding
W150A	0.09 ± 0.06	Ectoine binding
R280A	0.31 ± 0.14	Ectoine binding
R139A	5.01 ± 0.15	Dimerization
E140A	4.93 ± 0.20	Dimerization
R139A/E140A	4.72 ± 0.31	Dimerization

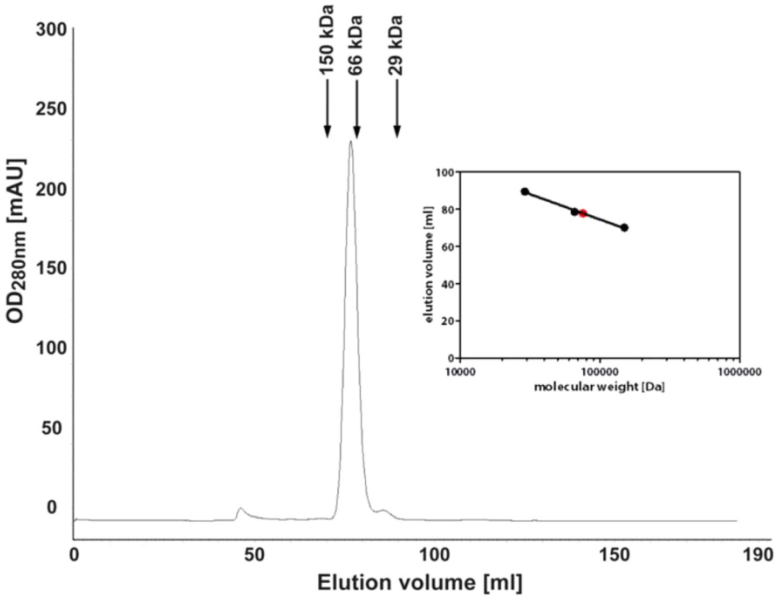


FIGURE 6. **Size-exclusion chromatography of the *Sa*EctD double mutant (R139A/Q140A).** The purified mutant protein (5 μg) was loaded onto a HiLoad 16/600 Superdex 200pg column (GE Healthcare) equilibrated and run in a 20 mM TES buffer, pH 7.5, containing 150 mM NaCl. The column run was performed at 15 °C. The size-exclusion column was standardized with a gel filtration markers kit (Sigma). The elution profile of the mutant EctD protein is shown, and *arrows* indicate that of relevant marker proteins. The *inset* shows a plot of the molecular mass (given in Da) of marker proteins (*black dots*) and of the EctD (R139A/Q140A) variant (*red dot*) *versus* the elution volume (given in ml).

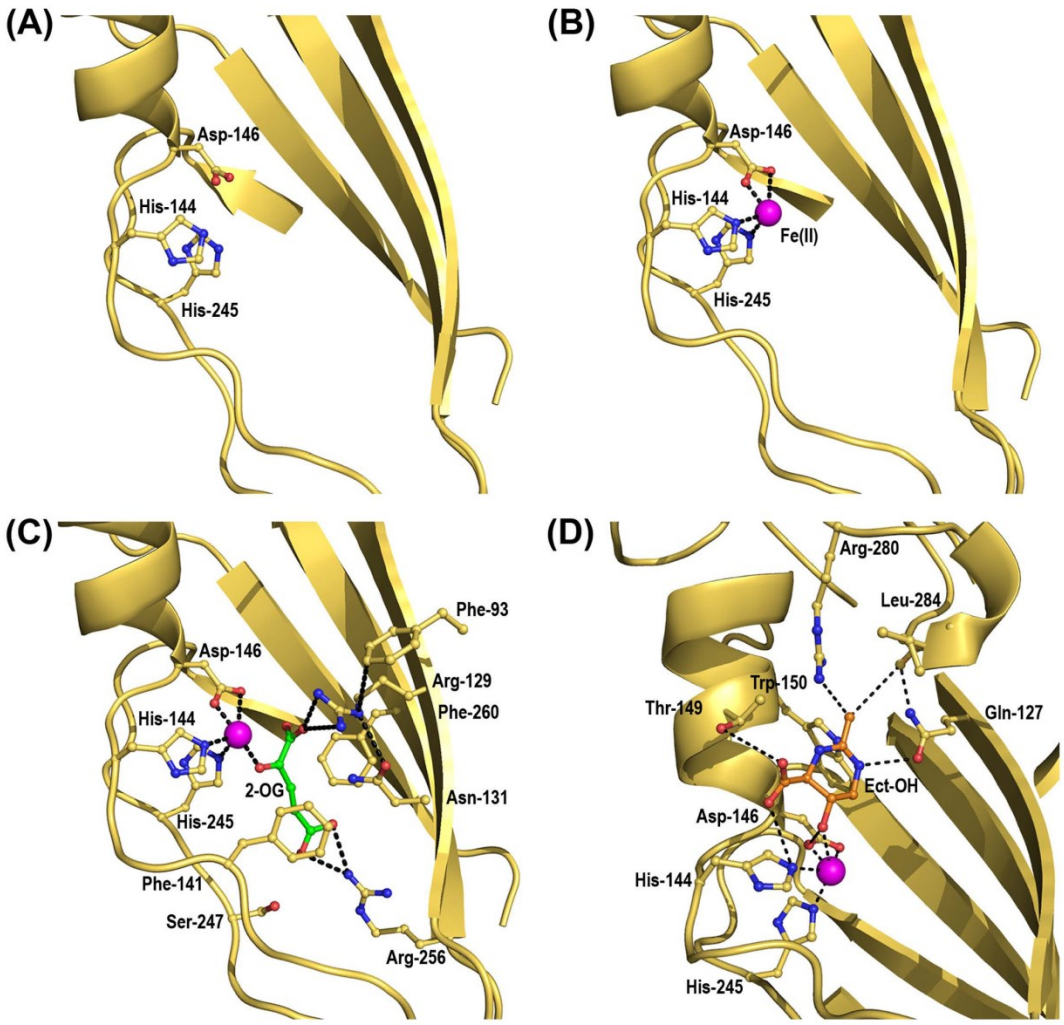


FIGURE 7. **Crystal structures of the *Sa*EctD-*Strep*-tag-II protein in its apo- and iron-bound forms and in complex with iron, 2-oxoglutarate, and 5-hydroxyectoine.** *Ribbon* representation of the catalytic core of the iron-free (PDB code 4HMR) (*A*) and iron-bound (PDB code 4MHU) (*B*) forms of the *S. alaskensis* ectoine hydroxylase. The ribbon representations of the catalytic core of the ectoine hydroxylase shown in *C* and *D* are both based on a crystal structure that contained the iron catalyst, the 2-oxoglutarate co-substrate, and the product of the EctD-catalyzed enzyme reaction, 5-hydroxyectoine (PDB code 4Q5O). In the structure shown in *C*, only the binding of the iron catalyst and the 2-oxoglutarate (*2-OG*) molecules are shown, and the structure shown in *D* highlights only the binding of the iron catalyst and the captured 5-hydroxyectoine (*Ect-OH*) molecule. Amino acids involved in ligand binding are represented as *sticks*.

Architecture of the Iron-binding Site—Because the overproduction conditions perfected for the large scale isolation of the ectoine hydroxylase in *E. coli* (see above) foster the synthesis of iron-free recombinant *Sa*EctD-*Strep*-tag-II protein, we were able to obtain *Sa*EctD crystal structures that either possessed or lacked the iron catalyst (Fig. 7, *A* and *B*). This was accomplished by adding or leaving out $Fe(II)Cl_2$ in the crystallization trials.

The iron ligand is bound in the *Sa*EctD protein via interaction with two histidine side chains, His-144 and His-245, and that of Asp-146 (Fig. 7*B*). These residues form a structurally well conserved H*X*(D/E) . . . H motif, the so-called 2-His-1-carboxylate facial triad, forming a type of mononuclear iron center found in many members of the dioxygenase superfamily (37–40, 64). The three iron-binding residues are fully conserved among a group of 433 aligned EctD-type proteins (22). Two of the iron-coordinating residues (His-144 and Asp-146) in the *Sa*EctD protein are part of a 17-amino acid residue region that is strictly conserved in ectoine hydroxylases (22, 27, 44). This EctD signature sequence spans an extended α-helix and a linked short β-sheet lining one side of the DSBH/cupin fold (Fig. 4*A*). We compared the structural arrangement of this region in the apo- and iron-bound forms of *Sa*EctD (Fig. 7, *A*

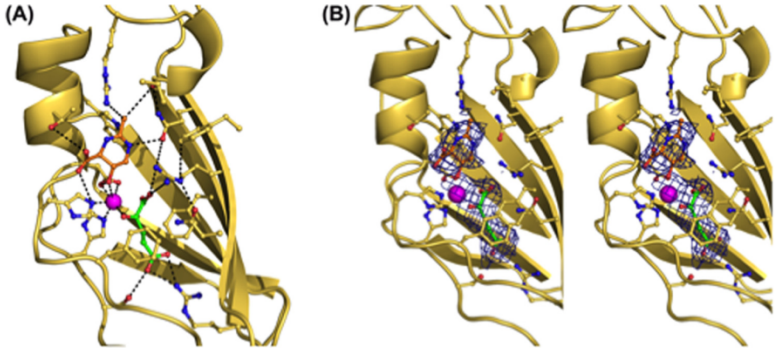


FIGURE 8. **Architecture of the active site of the *S. alaskensis* ectoine hydroxylase.** The crystal structure shown in *A* highlights the network of interactions allowing the binding of the iron catalyst and of the 2-oxoglutarate co-substrate and the EctD-catalyzed enzyme reaction product, 5-hydroxyectoine. Amino acids involved in ligand binding are represented as *sticks*. In (*B*) the electron density of an omit map (calculated by the Phenix software) of the 2-oxoglutarate and 5-hydroxyectoine molecules captured by the active site of *Sa*EctD is shown as stereo figure. The 2Fo-Fc map was contoured at 1.0 sigma. Both figures are based on data deposited in the PDB under accession code 4Q5O.

and *B*) and observed no significant differences between these two structures, thus reinforcing our conclusion (22) that the capturing of the iron ligand does not induce significant structural changes in the ectoine hydroxylase.

Architecture of the 2-Oxoglutarate-binding Site—In the crystals that we obtained by adding $Fe(II)Cl_2$, the co-substrate 2-oxoglutarate, and the reaction product 5-hydroxyectoine, the densities for the 2-oxoglutarate and 5-hydroxyectoine ligands were clearly visible. We first describe the coordination of the 2-oxoglutarate co-substrate and subsequently that of 5-hydroxyectoine by the EctD protein.

The 2-oxoacid group of 2-oxoglutarate is placed in close vicinity to the iron catalyst in the active site (Fig. 7*C*). Interactions with the side chains of Arg-256, Ser-247, and Phe-141 are key for 2-oxoglutarate binding (Fig. 7*C*). Furthermore, stabilizing interactions of this ligand are formed with the side chain of Arg-129. The precise configuration of the Arg-129 side chain is in turn stabilized through interactions with the side chains of Phe-93 and Asn-131 (Fig. 7*C*). Arg-129 and Asn-131 are part of the DSBH core; however, Phe-93 is positioned in a β-sheet flanking this fold. The Arg-129/Phe-93 and Arg-129/Asn-131 interactions thereby stabilize the conformation of the *Sa*EctD protein as a whole. Loss of these interactions would probably lay the active site of *Sa*EctD more open, possibly rendering it unable to bind any of its substrates in a coordinated fashion. The 2-oxoglutarate ligand within the active site is further stabilized by Phe-260, although this interaction appears to be not very strong because the distance between 2-oxoglutarate and the Phe side chain is almost 4 Å (Fig. 7*C*).

The architecture of the 2-oxoglutarate-binding site of the *Sa*EctD protein is similar to that found in crystal structures of many non-heme containing iron(II) and 2-oxoglutarate-dependent dioxygenases (38–40). It also corresponds closely to that suggested by structural comparison and modeling studies of the ligand-free *Vs*EctD protein (43, 44).

Architecture of the 5-Hydroxyectoine-binding Site—The 5-hydroxyectoine molecule is bound within the active site of *Sa*EctD slightly above the three residues (His-144, His-245, and Asp-146) forming the 2-His-1-carboxylate facial triad and in such a way that the hydroxy group at C5 in the pyrimidine ring points toward the iron catalyst (Fig. 7*D*). It is also positioned in close vicinity to the 2-oxoglutarate co-substrate (Fig. 8*A*). Hence, the spatial orientation of the 5-hydroxyectoine molecule is such as one would expect after the EctD enzyme has completed the hydroxylation of its substrate ectoine (27, 33, 43). The electron densities of the iron catalyst, the 2-oxoglutarate co-substrate, and the reaction product 5-hydroxyectoine are reasonably well resolved in the 2.56 Å *Sa*EctD crystal structure (Fig. 8*B*).

Interactions of the EctD protein with the ring structure of 5-hydroxyectoine and its methyl and carboxyl groups position and stabilize this molecule within the active site (Fig. 7*D*). The side chains of Gln-127, Trp-150, and Arg-280 are directly involved in these interactions; a hydrogen bond to the backbone of Leu-284 and a weak interaction with the side chain of Thr-149 (distance of 3.9–4.0 Å) provide additional stability for the binding of the ligand (Fig. 7*D*). Arg-129 interacts with the side chain of Gln-127 and therefore provides further stability (Fig. 5*A*). In addition, Trp-150, a residue that belongs to the ectoine hydroxylase signature sequence (22, 27, 44), makes contact with the pyrimidine ring of 5-hydroxyectoine (Fig. 7*D*). This residue seems to play a special role in ectoine/5-hydroxyectoine binding because its conservative replacement with either Phe or Tyr residues in the *Vs*EctD enzyme yielded catalytically inactive protein variants (43). This finding suggests that the interactions with the second aromatic ring structure of Trp are important for ligand stabilization. The carboxyl group of 5-hydroxyectoine is coordinated through interactions with Thr-149, a residue of the signature sequence motif, and His-144, which is part of the iron-coordinating 2-His-1-Asp facial triad motif (Fig. 7*D*).

Molecular dynamics simulations were recently used to glean information about the possible location of the ectoine substrate within the crystal structure of the *Vs*EctD protein (43). Despite the moderate affinity of this enzyme for its substrate ectoine (K_m = 5.9 ± 0.3 mM) (22), this modeling approach positioned the ectoine molecule to the same ligand-binding pocket that was revealed through the crystallographic analysis of the

*Sa*EctD protein presented in this study (Fig. 7*D*). However, relative to the spatial orientation of the 5-hydroxyectoine ligand found in the experimentally determined *Sa*EctD crystal structure (Figs. 7*D* and 8*A*), the ectoine molecule is flipped by 180° around its vertical axis when compared with its position in the *in silico* generated model, and it is also slightly tilted (43). In the *Sa*EctD crystal structure, the carboxylate of the 5-hydroxyectoine molecule is stabilized via hydrogen bonds (Figs. 7*D* and 8*A*), although such stabilizing interactions are missing in the *Vs*EctD *in silico* model (43). Interactions with the carboxylate of either ectoine or 5-hydroxyectoine are important for high affinity ligand binding by solute receptor proteins operating in conjunction with either ABC or TRAP transport systems for these compounds (65–67). We therefore surmise that the *Sa*EctD crystal structure in complex with 5-hydroxyectoine (Fig. 8*A*) also reflects the configuration and spatial position of the ectoine substrate when it is captured in the active site by the ectoine hydroxylase prior to catalysis.

Structure-guided Mutagenesis of 5-Hydroxyectoine-contacting Residues—The binding of the iron catalyst and the 2-oxoglutarate co-substrate in non-heme-containing dioxygenases is mediated by structurally well conserved motifs (37–40), and these are also present in the *Vs*EctD (22, 44) and *Sa*EctD proteins (Fig. 7, *B* and *C*). The corresponding residues in the *Vs*EctD protein (Fig. 3) have already been targeted by site-directed mutagenesis and inevitably yielded mutant EctD proteins that were either strongly impaired in their enzyme activity or were catalytically inactive (43). In view of the fact that the corresponding iron- and 2-oxoglutarate-coordinating residues (Fig. 4, *B* and *C*) within the EctD enzyme family are practically completely conserved within a group of 433 inspected EctD-type proteins (22, 43), we have not repeated such experiments with the *Sa*EctD protein. Instead, we have focused on those residues that contact the 5-hydroxyectoine molecule within the ligand-binding site of *Sa*EctD (Fig. 7*D*).

Residues Gln-127, Thr-149, Trp-150, and Arg-280 (Fig. 7*D*) were individually substituted by an Ala residue; Leu-284 was not changed because it makes a backbone contact to 5-hydroxyectoine. Likewise, the 5-hydroxyectoine-contacting residues His-144 and Asp-146 were also not targeted by site-directed mutagenesis because their side chains are not only involved in 5-hydroxyectoine binding but are critical for binding of the iron catalyst as well (Fig. 7, *B* and *D*). The individual substitution of Gln-127, Thr-149, Trp-150, and Arg-280 by Ala residues yielded EctD variants that were either strongly impaired in their enzyme function (Thr-149/Ala and Arg-280/Ala), or almost completely catalytically inactive (Gln-127/Ala and Trp-150/Ala) (Table 3). Fully consistent with these results are data from a mutagenesis study of the *Vs*EctD protein in which the two 5-hydroxyectoine-contacting residues corresponding to Gln-127 and Trp-150 in the *Sa*EctD protein (Fig. 3) have been studied; their replacement yielded enzymatically inactive variants of EctD as well (43). Given the combined data from the site-directed mutagenesis studies of the *Sa*EctD (Table 3) and the *Vs*EctD (43) proteins, it is not unexpected that each residue involved in the binding of 5-hydroxyectoine is completely conserved in an alignment of 433 microbial EctD-type proteins (22, 43).

DISCUSSION

Members of the superfamily of non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases share similar overall structures and enzyme mechanisms, but they can catalyze a diverse set of oxidation reactions (37–41). The ectoine hydroxylase (EctD) belongs to this superfamily and mediates the stereo-specific hydroxylation of the compatible solute ectoine to form 5-hydroxyectoine (27, 68), which thereby gains novel stress-protective and function-preserving properties (11–15, 35, 36). The crystal structures of EctD from the extremophilic bacteria *V. salexigens* (44) and *S. alaskensis* (this study) are most closely related to the human hydroxylases PhyH and PHYD1A (55, 56) and to the microbial halogenases SyrB2, CytC3, and CurA (57–59). It is noteworthy that the substrates of the mentioned halogenases are all tethered via a thioester to the phosphopantetheine arms of acyl carrier proteins of polyketide synthases and nonribosomal peptide synthetases that function within assembly lines directing the synthesis of antibiotics and phytotoxins (57–59). In contrast, ectoine hydroxylases use a freely diffusible small molecule as their substrate (22, 26, 27).

Various quaternary conformations are found for enzymes belonging to the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily (37–39, 62). Conventional size-exclusion chromatography (22) and HPLC-MALS analysis (Fig. 2) strongly suggest that ectoine hydroxylases obtained from various producer organisms are all dimers in solution. In the homodimer of the *Sa*EctD crystal structure, the two monomers are oriented in a head to tail fashion and interact through loop regions (Fig. 4).

The crystallographic analysis of the *Sa*EctD protein in complex with all its ligands (Fig. 8*A*) revealed that three residues contribute to iron binding, seven are involved in the binding of 2-oxoglutarate, and seven mediate the binding of 5-hydroxyectoine. Of these 15 residues, His-144 and Asp-146 are simultaneously involved in the binding of the iron catalyst and of 5-hydroxyectoine (Fig. 8*A*).

When one views the overall shape and surface structure of the crystallized EctD proteins as a whole, a deep cavity is immediately apparent in each of the monomers (Fig. 4*B*). The iron ligand, the co-substrate 2-oxoglutarate, and the EctD reaction product 5-hydroxyectoine (and by inference also the substrate ectoine) are all located in this cavity (Fig. 4*B*). As expected, all residues that we deem to be of functional importance for ligand binding and catalysis protrude into this cavity of the EctD protein (Figs. 4*B* and 8*A*).

Multiple cation/π interactions between the delocalized positive charges of the ectoine and 5-hydroxyectoine molecules (Fig. 1) and the side chains of aromatic residues of extracellular ectoine/5-hydroxyectoine-specific solute receptor proteins (EhuB, UehA, and TeaA) that operate in conjunction with either ABC or TRAP transport systems (65–67) are key determinants for high affinity binding of these tetrahydropyrimidines. These ligand-binding proteins possess K_d values in the low micromolar range because they need to scavenge ectoines from scarce environmental resources for use either as osmostress protectants or as nutrients (65–67). Ectoine hydroxy-

lases, however, have K_m values in the low millimolar range (between 5 and 10 mM) (22, 26, 27), and this property might be the reason why one typically finds considerable amounts of ectoine in osmotically stressed microbial cells before 5-hydroxyectoine production sets in (26, 27, 69). In contrast to the aforementioned solute receptor proteins for ectoines (65–67), cation/π interactions do not contribute to 5-hydroxyectoine binding (and by inference to the substrate ectoine as well) by the EctD enzyme (Fig. 5*A*). This might contribute to the rather modest affinity of ectoine hydroxylases for their substrate ectoine (22, 26, 27).

Alignments of the amino acid sequences of ectoine hydroxylases have revealed a signature sequence motif of 17 amino acids in length (FXWHSDFETWHXEDG(M/L)P) (22, 27, 43, 44) that is, with a few minor exceptions, strictly conserved in 433 EctD-type proteins inspected by us (22). This string of amino acids forms an extended α-helix that is connected to a short β-strand; it structures one side of the EctD cupin barrel (Fig. 4*A*). Not only is the signature sequence region structurally important for the overall fold of the EctD protein but the crystallographic data presented here show that it also contains five residues that contribute in multiple ways to the binding of all three ligands of the ectoine hydroxylase. In the *Sa*EctD protein, these residues include the side chains of His-144 and Asp-146 that are involved in iron binding, Phe-141, which is involved in 2-oxoglutarate binding, and the 5-hydroxyectoine (ectoine)-contacting amino acids His-144, Asp-146, Thr-149, and Trp-150 (Fig. 8).

When one views the *Sa*EctD active site as a whole, an intricate network of interactions between the iron catalyst, the co-substrate 2-oxoglutarate and the 5-hydroxyectoine (and by inference ectoine) and their corresponding binding partners in the catalytic core of the ectoine hydroxylase becomes apparent (Fig. 8*A*). This network of interactions mediates the precise positioning of the ectoine substrate in such way that the C5 carbon atom in its pyrimidine ring can be hydroxylated in a stereo-specific manner. The ectoine hydroxylase catalyzes this enzyme reaction with high accuracy, both *in vivo* (33) and *in vitro* (27).

The EctD enzyme is not only of interest from an ecophysiological and biotechnological point of view (9, 10), it also has the potential for use in chemical biology with respect to the generation of chemically modified ectoine molecules. Synthetic ectoine derivatives with either expanded or reduced ring sizes have already been reported (70, 71), and these might lead to interesting novel biotechnological applications or medical uses (9, 10, 17). The architecture of the catalytic core of the ectoine hydroxylase (Fig. 7 and 8) is probably flexible enough to allow chemical modifications of these existing synthetic molecules. An EctD-mediated hydroxylation might endow them with novel stress-protective and structure-preserving functions in the same way that it allows 5-hydroxyectoine to function strikingly different from ectoine in alleviating desiccation stress (36).

The structural insight that we provide here for the active site of the EctD protein (Figs. 7 and 8) might therefore aid the rational chemical design of new ectoine derivatives that then could be hydroxylated either *in vivo* or *in vitro* by the EctD enzyme. Furthermore, our crystallographic study might also inspire experiments to improve the moderate catalytic efficiencies of ectoine hydroxylases (22) through targeted mutagenesis or *in vivo* evolution experiments to enhance their industrial use.

Acknowledgments—We thank Jochen Sohn for excellent technical assistance in the overproduction and purification of the EctD protein and Georg Lentzen and Irina Bagyan (bitop AG) for their kind gifts of ectoine and 5-hydroxyectoine. We thank the staff of the P13 and P14 beamlines at EMBL, Hamburg, Germany, for their kind support during crystal screening. We acknowledge the European Synchrotron Radiation Facility for provision of synchrotron radiation facilities (ID23eh2). We greatly appreciate the expert help of Vickie Koogle in the language editing of our manuscript and thank Lutz Schmitt for helpful discussions and the kind support for this project. E. B. greatly valued the hospitality of Tom Silhavy during a sabbatical at the Department of Molecular Biology, Princeton University (Princeton, NJ). N. W. and E. B. are very grateful to Rolf Thauer for his continued support.

REFERENCES

1. Yancey, P. H. (2005) Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *J. Exp. Biol.* **208,** 2819–2830
2. Kempf, B., and Bremer, E. (1998) Uptake and synthesis of compatible solutes as microbial stress responses to high osmolality environments. *Arch. Microbiol.* **170,** 319–330
3. Bremer, E., and Krämer, R. (2000) in *Bacterial Stress Responses* (Storz, G., and Hengge-Aronis, R., eds) pp. 79–97, American Society for Microbiology, Washington, D. C.
4. Street, T. O., Bolen, D. W., and Rose, G. D. (2006) A molecular mechanism for osmolyte-induced protein stability. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 13997–14002
5. Ignatova, Z., and Gierasch, L. M. (2006) Inhibition of protein aggregation *in vitro* and *in vivo* by a natural osmoprotectant. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 13357–13361
6. Bourot, S., Sire, O., Trautwetter, A., Touzé, T., Wu, L. F., Blanco, C., and Bernard, T. (2000) Glycine betaine-assisted protein folding in a *lysA* mutant of *Escherichia coli*. *J. Biol. Chem.* **275,** 1050–1056
7. Diamant, S., Eliahu, N., Rosenthal, D., and Goloubinoff, P. (2001) Chemical chaperones regulate molecular chaperones *in vitro* and in cells under combined salt and heat stresses. *J. Biol. Chem.* **276,** 39586–39591
8. Galinski, E. A., Pfeiffer, H. P., and Trüper, H. G. (1985) 1,4,5,6-Tetrahydro-2-methyl-4-pyrimidinecarboxylic acid. A novel cyclic amino acid from halophilic phototrophic bacteria of the genus *Ectothiorhodospira*. *Eur. J. Biochem.* **149,** 135–139
9. Lentzen, G., and Schwarz, T. (2006) Extremolytes: Natural compounds from extremophiles for versatile applications. *Appl. Microbiol. Biotechnol.* **72,** 623–634
10. Pastor, J. M., Salvador, M., Argandoña, M., Bernal, V., Reina-Bueno, M., Csonka, L. N., Iborra, J. L., Vargas, C., Nieto, J. J., and Cánovas, M. (2010) Ectoines in cell stress protection: uses and biotechnological production. *Biotechnol. Adv.* **28,** 782–801
11. Lippert, K., and Galinski, E. A. (1992) Enzyme stabilization by ectoine-type compatible solutes: protection against heating, freezing and drying. *Appl. Microbiol. Biotechnol.* **37,** 61–65
12. Knapp, S., Ladenstein, R., and Galinski, E. A. (1999) Extrinsic protein stabilization by the naturally occurring osmolytes β-hydroxyectoine and betaine. *Extremophiles* **3,** 191–198
13. Borges, N., Ramos, A., Raven, N. D., Sharp, R. J., and Santos, H. (2002) Comparative study of the thermostabilizing properties of mannosylglycerate and other compatible solutes on model enzymes. *Extremophiles* **6,** 209–216
14. Van-Thuoc, D., Hashim, S. O., Hatti-Kaul, R., and Mamo, G. (2013) Ecto-

ine-mediated protection of enzyme from the effect of pH and temperature stress: a study using *Bacillus halodurans* xylanase as a model. *Appl. Microbiol. Biotechnol.* **97,** 6271–6278
15. Manzanera, M., García de Castro, A., Tøndervik, A., Rayner-Brandes, M., Strøm, A. R., and Tunnacliffe, A. (2002) Hydroxyectoine is superior to trehalose for anhydrobiotic engineering of *Pseudomonas putida* KT2440. *Appl. Environ. Microbiol.* **68,** 4328–4333
16. Schwibbert, K., Marin-Sanguino, A., Bagyan, I., Heidrich, G., Lentzen, G., Seitz, H., Rampp, M., Schuster, S. C., Klenk, H. P., Pfeiffer, F., Oesterhelt, D., and Kunte, H. J. (2011) A blueprint of ectoine metabolism from the genome of the industrial producer *Halomonas elongata* DSM 2581 T. *Environ. Microbiol.* **13,** 1973–1994
17. Graf, R., Anzali, S., Buenger, J., Pfluecker, F., and Driller, H. (2008) The multifunctional role of ectoine as a natural cell protectant. *Clin. Dermatol.* **26,** 326–333
18. Becker, J., Schäfer, R., Kohlstedt, M., Harder, B. J., Borchert, N. S., Stöveken, N., Bremer, E., and Wittmann, C. (2013) Systems metabolic engineering of *Corynebacterium glutamicum* for production of the chemical chaperone ectoine. *Microb. Cell Fact.* **12,** 110
19. Marini, A., Reinelt, K., Krutmann, J., and Bilstein, A. (2014) Ectoine-containing cream in the treatment of mild to moderate atopic dermatitis: a randomised, comparator-controlled, intra-individual double-blind, multi-center trial. *Skin Pharmacol. Physiol.* **27,** 57–65
20. Abdel-Aziz, H., Wadie, W., Abdallah, D. M., Lentzen, G., and Khayyal, M. T. (2013) Novel effects of ectoine, a bacteria-derived natural tetrahydropyrimidine, in experimental colitis. *Phytomedicine* **20,** 585–591
21. Sun, H., Glasmacher, B., and Hofmann, N. (2012) Compatible solutes improve cryopreservation of human endothelial cells. *Cryo Lett.* **33,** 485–493
22. Widderich, N., Höppner, A., Pittelkow, M., Heider, J., Smits, S. H., and Bremer, E. (2014) Biochemical properties of ectoine hydroxylases from extremophiles and their wider taxonomic distribution among microorganisms. *PLoS One* **9,** e93809
23. Louis, P., and Galinski, E. A. (1997) Characterization of genes for the biosynthesis of the compatible solute ectoine from *Marinococcus halophilus* and osmoregulated expression in *Escherichia coli*. *Microbiology* **143,** 1141–1149
24. Kuhlmann, A. U., and Bremer, E. (2002) Osmotically regulated synthesis of the compatible solute ectoine in *Bacillus pasteurii* and related *Bacillus* spp. *Appl. Environ. Microbiol.* **68,** 772–783
25. Kuhlmann, A. U., Bursy, J., Gimpel, S., Hoffmann, T., and Bremer, E. (2008) Synthesis of the compatible solute ectoine in *Virgibacillus pantothenticus* is triggered by high salinity and low growth temperature. *Appl. Environ. Microbiol.* **74,** 4560–4563
26. Bursy, J., Kuhlmann, A. U., Pittelkow, M., Hartmann, H., Jebbar, M., Pierik, A. J., and Bremer, E. (2008) Synthesis and uptake of the compatible solutes ectoine and 5-hydroxyectoine by *Streptomyces coelicolor* A3(2) in response to salt and heat stresses. *Appl. Environ. Microbiol.* **74,** 7286–7296
27. Bursy, J., Pierik, A. J., Pica, N., and Bremer, E. (2007) Osmotically induced synthesis of the compatible solute hydroxyectoine is mediated by an evolutionarily conserved ectoine hydroxylase. *J. Biol. Chem.* **282,** 31147–31155
28. Saum, S. H., and Müller, V. (2008) Growth phase-dependent switch in osmolyte strategy in a moderate halophile: ectoine is a minor osmolyte but major stationary phase solute in *Halobacillus halophilus*. *Environ. Microbiol.* **10,** 716–726
29. Mustakhimov, I. I., Reshetnikov, A. S., Glukhov, A. S., Khmelenina, V. N., Kalyuzhnaya, M. G., and Trotsenko, Y. A. (2010) Identification and characterization of EctR1, a new transcriptional regulator of the ectoine biosynthesis genes in the halotolerant methanotroph *Methylomicrobium alcaliphilum* 20Z. *J. Bacteriol.* **192,** 410–417
30. García-Estepa, R., Argandoña, M., Reina-Bueno, M., Capote, N., Iglesias-Guerra, F., Nieto, J. J., and Vargas, C. (2006) The *ectD* gene, which is involved in the synthesis of the compatible solute hydroxyectoine, is essential for thermoprotection of the halophilic bacterium *Chromohalobacter salexigens*. *J. Bacteriol.* **188,** 3774–3784
31. Peters, P., Galinski, E. A., and Trüper, H. G. (1990) The biosyntheis of ectoine. *FEMS Microbiol. Lett.* **71,** 157–162
32. Ono, H., Sawada, K., Khunajakr, N., Tao, T., Yamamoto, M., Hiramoto, M., Shinmyo, A., Takano, M., and Murooka, Y. (1999) Characterization of biosynthetic enzymes for ectoine as a compatible solute in a moderately halophilic eubacterium, *Halomonas elongata*. *J. Bacteriol.* **181,** 91–99
33. Inbar, L., and Lapidot, A. (1988) The structure and biosynthesis of new tetrahydropyrimidine derivatives in actinomycin D producer *Streptomyces parvulus*. Use of ^{13}C- and ^{15}N-labeled L-glutamate and ^{13}C and ^{15}N NMR spectroscopy. *J. Biol. Chem.* **263,** 16014–16022
34. Prabhu, J., Schauwecker, F., Grammel, N., Keller, U., and Bernhard, M. (2004) Functional expression of the ectoine hydroxylase gene (*thpD*) from *Streptomyces chrysomallus* in *Halomonas elongata*. *Appl. Environ. Microbiol.* **70,** 3130–3132
35. Kurz, M. (2008) Compatible solute influence on nucleic acids: many questions but few answers. *Saline Systems* **4,** 6
36. Tanne, C., Golovina, E. A., Hoekstra, F. A., Meffert, A., and Galinski, E. A. (2014) Glass-forming property of hydroxyectoine is the cause of its superior function as a dessication protectant. *Front. Microbiol.* **5,** 150
37. Hausinger, R. P. (2004) FeII/α-ketoglutarate-dependent hydroxylases and related enzymes. *Crit. Rev. Biochem. Mol. Biol.* **39,** 21–68
38. Aik, W., McDonough, M. A., Thalhammer, A., Chowdhury, R., and Schofield, C. J. (2012) Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. *Curr. Opin. Struct. Biol.* **22,** 691–700
39. Hangasky, J. A., Taabazuing, C. Y., Valliere, M. A., and Knapp, M. J. (2013) Imposing function down a (cupin)-barrel: secondary structure and metal stereochemistry in the αKG-dependent oxygenases. *Metallomics* **5,** 287–301
40. Kundu, S. (2012) Distribution and prediction of catalytic domains in 2-oxoglutarate-dependent dioxygenases. *BMC Res. Notes* **5,** 410
41. Wong, S. D., Srnec, M., Matthews, M. L., Liu, L. V., Kwak, Y., Park, K., Bell, C. B., 3rd, Alp, E. E., Zhao, J., Yoda, Y., Kitao, S., Seto, M., Krebs, C., Bollinger, J. M., Jr., and Solomon, E. I. (2013) Elucidation of the Fe(IV)=O intermediate in the catalytic cycle of the halogenase SyrB2. *Nature* **499,** 320–323
42. Grzyska, P. K., Appelman, E. H., Hausinger, R. P., and Proshlyakov, D. A. (2010) Insight into the mechanism of an iron dioxygenase by resolution of steps following the FeIV=HO species. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 3982–3987
43. Widderich, N., Pittelkow, M., Höppner, A., Mulnaes, D., Buckel, W., Gohlke, H., Smits, S. H., and Bremer, E. (2014) Molecular dynamics simulations and structure-guided mutagenesis provide insight into the architecture of the catalytic core of the ectoine hydroxylase. *J. Mol. Biol.* **426,** 586–600
44. Reuter, K., Pittelkow, M., Bursy, J., Heine, A., Craan, T., and Bremer, E. (2010) Synthesis of 5-hydroxyectoine from ectoine: crystal structure of the non-heme iron(II) and 2-oxoglutarate-dependent dioxygenase EctD. *PLoS One* **5,** e10647
45. Ting, L., Williams, T. J., Cowley, M. J., Lauro, F. M., Guilhaus, M., Raftery, M. J., and Cavicchioli, R. (2010) Cold adaptation in the marine bacterium, *Sphingopyxis alaskensis*, assessed using quantitative proteomics. *Environ. Microbiol.* **12,** 2658–2676
46. Hoeppner, A., Widderich, N., Bremer, E., and Smits, S. H. J. (2014) Overexpression, crystallization and preliminary X-ray crystallographic analysis of the ectoine hydroxylase from *Sphingopyxis alaskensis*. *Acta Crystallogr. F Struct. Biol. Crystalliz. Commun.* **70,** 493–496
47. Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
48. Lovenberg, W., Buchanan, B. B., and Rabinowitz, J. C. (1963) Studies on the chemical nature of *Clostridial ferredoxin*. *J. Biol. Chem.* **238,** 3899–3913
49. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132
50. Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 133–144
51. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674
52. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 486–501
53. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of

macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53,** 240–255
54. Hoeppner, A., Schmitt, L., and Smits, S. H. (2013) in *Advances Topics on Crystal Growth* (Ferreira, S., ed) pp. 3–44, InTech Europe, Janeza Trdine 9, 51000 Rijeka, Croatia
55. McDonough, M. A., Kavanagh, K. L., Butler, D., Searls, T., Oppermann, U., and Schofield, C. J. (2005) Structure of human phytanoyl-CoA 2-hydroxylase identifies molecular mechanisms of Refsum disease. *J. Biol. Chem.* **280,** 41101–41110
56. Zhang, Z., Kochan, G. T., Ng, S. S., Kavanagh, K. L., Oppermann, U., Schofield, C. J., and McDonough, M. A. (2011) Crystal structure of PHYHD1A, a 2OG oxygenase related to phytanoyl-CoA hydroxylase. *Biochem. Biophys. Res. Commun.* **408,** 553–558
57. Blasiak, L. C., Vaillancourt, F. H., Walsh, C. T., and Drennan, C. L. (2006) Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis. *Nature* **440,** 368–371
58. Wong, C., Fujimori, D. G., Walsh, C. T., and Drennan, C. L. (2009) Structural analysis of an open active site conformation of nonheme iron halogenase CytC3. *J. Am. Chem. Soc.* **131,** 4872–4879
59. Khare, D., Wang, B., Gu, L., Razelun, J., Sherman, D. H., Gerwick, W. H., Håkansson, K., and Smith, J. L. (2010) Conformational switch triggered by α-ketoglutarate in a halogenase of curacin A biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 14099–14104
60. Strieker, M., Kopp, F., Mahlert, C., Essen, L. O., and Marahiel, M. A. (2007) Mechanistic and structural basis of stereospecific Cβ-hydroxylation in calcium-dependent antibiotic, a daptomycin-type lipopeptide. *ACS Chem. Biol.* **2,** 187–196
61. Helmetag, V., Samel, S. A., Thomas, M. G., Marahiel, M. A., and Essen, L. O. (2009) Structural basis for the erythro-stereospecificity of the L-arginine oxygenase VioC in viomycin biosynthesis. *FEBS J.* **276,** 3669–3682
62. Knauer, S. H., Hartl-Spiegelhauer, O., Schwarzinger, S., Hänzelmann, P., and Dobbek, H. (2012) The Fe(II)/α-ketoglutarate-dependent taurine dioxygenases from *Pseudomonas putida* and *Escherichia coli* are tetramers. *FEBS J.* **279,** 816–831
63. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372,** 774–797
64. Straganz, G. D., and Nidetzky, B. (2006) Variations of the 2-His-1-carboxylate theme in mononuclear non-heme FeII oxygenases. *Chembiochem* **10,** 1536–1548
65. Lecher, J., Pittelkow, M., Zobel, S., Bursy, J., Bönig, T., Smits, S. H., Schmitt, L., and Bremer, E. (2009) The crystal structure of UehA in complex with ectoine-A comparison with other TRAP-T binding proteins. *J. Mol. Biol.* **389,** 58–73
66. Hanekop, N., Höing, M., Sohn-Bösser, L., Jebbar, M., Schmitt, L., and Bremer, E. (2007) Crystal structure of the ligand-binding protein EhuB from *Sinorhizobium meliloti* reveals substrate recognition of the compatible solutes ectoine and hydroxyectoine. *J. Mol. Biol.* **374,** 1237–1250
67. Kuhlmann, S. I., Terwisscha van Scheltinga, A. C., Bienert, R., Kunte, H. J., and Ziegler, C. (2008) 1.55 A structure of the ectoine binding protein TeaA of the osmoregulated TRAP-transporter TeaABC from *Halomonas elongata*. *Biochemistry* **47,** 9475–9485
68. Inbar, L., Frolow, F., and Lapidot, A. (1993) The conformation of new tetrahydropyrimidine derivatives in solution and in the crystal. *Eur. J. Biochem.* **214,** 897–906
69. Stöveken, N., Pittelkow, M., Sinner, T., Jensen, R. A., Heider, J., and Bremer, E. (2011) A specialized aspartokinase enhances the biosynthesis of the osmoprotectants ectoine and hydroxyectoine in *Pseudomonas stutzeri* A1501. *J. Bacteriol.* **193,** 4456–4468
70. Schnoor, M., Voss, P., Cullen, P., Böking, T., Galla, H. J., Galinski, E. A., and Lorkowski, S. (2004) Characterization of the synthetic compatible solute homoectoine as a potent PCR enhancer. *Biochem. Biophys. Res. Commun.* **322,** 867–872
71. Witt, E. M., Davies, N. W., and Galinski, E. A. (2011) Unexpected property of ectoine synthase and its application for synthesis of the engineered compatible solute ADPC. *Appl. Microbiol. Biotechnol.* **91,** 113–122

V. Publication 5

The fifth publication is entitled ‘Overproduction, crystallization and X-ray diffraction data analysis of ectoine synthase from the cold-adapted marine bacterium *Sphingopyxis alaskensis*’. It was published in ‘Acta Crystallographica Section F Structural Biology Communications’ (doi: 10.1107/S2053230X15011115).

Kobus, S., Widderich, N., Hoeppner, A., Bremer, E., and Smits, S.H. Overproduction, crystallization and X-ray diffraction data analysis of ectoine synthase from the cold-adapted marine bacterium *Sphingopyxis alaskensis*. *Acta Crystallogr F Struct Biol Commun*. (2015). 71: 1027-1032.

This paper describes a detailed expression and purification protocol for the ectoine synthase from the cold-adapted marine bacterium *Sphingopyxis alaskensis*. It further identifies EctC to form dimers in solution and deals with crystallization trials and preliminary X-ray diffraction data of the EctC enzyme.

Personal contribution:
I established production and purification protocols for the *Sa*EctC protein, performed the corresponding experiments and conducted size-exclusion chromatography. I further analyzed the experimental data (together with the other authors) and contributed to the preparation of Figures (together with Sander Smits). I also contributed to the writing of the manuscript together with my PhD supervisor Erhard Bremer and Sander Smits.

research communications

STRUCTURAL BIOLOGY COMMUNICATIONS
ISSN 2053-230X

Received 13 April 2015
Accepted 8 June 2015

Edited by M. S. Weiss, Helmholtz-Zentrum Berlin für Materialien und Energie, Germany

Keywords: compatible solute; osmostress protectant; chemical chaperone; enzyme; ectoine synthesis; cupin; X-ray analysis.

Supporting information: this article has supporting information at journals.iucr.org/f

Overproduction, crystallization and X-ray diffraction data analysis of ectoine synthase from the cold-adapted marine bacterium *Sphingopyxis alaskensis*

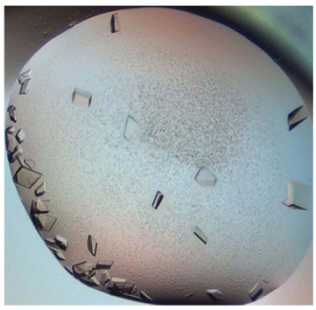
Stefanie Kobus,[a] Nils Widderich,[b] Astrid Hoeppner,[a] Erhard Bremer[b,c]* and Sander H. J. Smits[d]*

[a]Crystal Farm and X-ray Facility, Heinrich-Heine-University, Universitaetsstrasse 1, 40225 Düsseldorf, Germany, [b]Department of Biology, Laboratory for Microbiology, Philipps-University Marburg, Karl-von-Frisch Strasse 8, 35043 Marburg, Germany, [c]LOEWE Center for Synthetic Microbiology, Philipps-University Marburg, Hans-Meerwein-Strasse 6, 35043 Marburg, Germany, and [d]Institute of Biochemistry, Heinrich-Heine-University, Universitaetsstrasse 1, 40225 Düsseldorf, Germany. *Correspondence e-mail: bremer@staff.uni-marburg.de, sander.smits@hhu.de

Ectoine biosynthetic genes (*ectABC*) are widely distributed in bacteria. Microorganisms that carry them make copious amounts of ectoine as a cell protectant in response to high-osmolarity challenges. Ectoine synthase (EctC; EC 4.2.1.108) is the key enzyme for the production of this compatible solute and mediates the last step of ectoine biosynthesis. It catalyzes the ring closure of the cyclic ectoine molecule. A codon-optimized version of *ectC* from *Sphingopyxis alaskensis* (*Sa*) was used for overproduction of *Sa*EctC protein carrying a *Strep*-tag II peptide at its carboxy-terminus. The recombinant *Sa*EctC-*Strep*-tag II protein was purified to near-homogeneity from *Escherichia coli* cell extracts by affinity chromatography. Size-exclusion chromatography revealed that it is a dimer in solution. The *Sa*EctC-*Strep*-tag II protein was crystallized using the sitting-drop vapour-diffusion method and crystals that diffracted to 1.0 Å resolution were obtained.

1. Introduction

Compatible solutes are highly water-soluble organic osmolytes that are compliant with cellular biochemistry and physiology and can therefore be accumulated to exceedingly high intracellular levels. Microorganisms amass them, either through synthesis or uptake, to offset the detrimental consequences of a rise in the extracellular osmolarity on cellular hydration, turgor and growth (Kempf & Bremer, 1998). Ectoine [(4*S*)-2-methyl-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid; Fig. 1*a*; Galinski *et al.*, 1985] is such a compatible solute, and the genes (*ectABC*) encoding its three biosynthetic enzymes (Louis & Galinski, 1997; Ono *et al.*, 1999; Supplementary Fig. S1) are widely distributed among members of the Bacteria (Widderich *et al.*, 2014). In addition to their function as physiologically well tolerated osmotic and temperature stress protectants (Bursy *et al.*, 2008; García-Estepa *et al.*, 2006; Kuhlmann *et al.*, 2011), ectoine and its derivative 5-hydroxyectoine [(4*S*,5*S*)-2-methyl-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid; Höppner *et al.*, 2014; Supplementary Fig. S1] also serve as chemical chaperones since they can effectively preserve the functionality of biological macromolecules (Lippert & Galinski, 1992; Tanne *et al.*, 2014; Manzanera *et al.*, 2002). This has fostered a substantial interest in these compounds for their biotechnological exploitation as stabilizers of proteins, membranes and cells, for use in cosmetics and for potential medical applications (Pastor *et al.*, 2010; Graf *et al.*, 2008).



research communications

The biosynthesis of ectoine proceeds from L-aspartate-β-semialdehyde, a central hub in microbial amino-acid metabolism, and is mediated by the sequential reactions of L-2,4-diaminobutyrate transaminase (EctB; EC 2.6.1.76), 2,4-diaminobutyrate acetyltransferase (EctA; EC 2.3.1.178) and ectoine synthase (EctC; EC 4.2.1.108; Supplementary Fig. S1; Louis & Galinski, 1997; Ono *et al.*, 1999). The corresponding three enzymes from the highly salt-tolerant bacterium *Halomonas elongata* have already been biochemically characterized to some extent (Ono *et al.*, 1999). The crystal structure of EctA from *Bordetella parapertussis* in complex with its substrate L-2,4-diaminobutyrate (PDB entry 3d3s) has been determined within the framework of a structural genomics project. Likewise, the crystal structure of ectoine hydroxylase (EctD; EC 1.14.11) from *Sphingopyxis alaskensis* with all of its ligands and the reaction product 5-hydroxyectoine (PDB entry 4q5o) has also been reported (Höppner *et al.*, 2014).

Here, we focus on ectoine synthase, the key enzyme of the ectoine biosynthetic route (Ono *et al.*, 1999; Louis & Galinski, 1997) and a protein that can also be used as a diagnostic tool in database searches of microbial genome sequences to scan for potential microbial ectoine producers (Widderich *et al.*, 2014). A *BLAST* search of microbial genomes deposited in the NCBI database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) suggests that EctC belongs to the cupin superfamily, a large group of proteins with a conserved β-barrel structural scaffold that perform diverse enzymatic and non-enzymatic functions (Uberto & Moomaw, 2013). An initial biochemical characterization of the ectoine synthases from the extremophiles *H. elongata*, *Methylomicrobium alcaliphilum* and *Acidiphilium cryptum* has already been carried out (Ono *et al.*, 1999; Witt *et al.*, 2011; Moritz *et al.*, 2015; Reshetnikov *et al.*, 2006). Each of these enzymes catalyzes the cyclization of N_γ-acetyl-L-2,4-diaminobutyrate (Fig. 1*a*), the reaction product of EctA (Supplementary Fig. S1), to ectoine with the concomitant release of a water molecule (Fig. 1*a*). In a side reaction, EctC can also hydrolyze ectoine and some synthetic ectoine derivatives with altered ring sizes (Witt *et al.*, 2011; Moritz *et al.*, 2015). A further minor side reaction of ectoine synthase is the formation of the synthetic compatible solute 5-amino-3,4-dihydro-2*H*-pyrrole-2-carboxylate (ADPC) through the cyclic condensation of glutamine (Witt *et al.*, 2011).

Although progress has been made with respect to the biochemical characterization of the main activity of ectoine synthase, the cyclization of N_γ-acetyl-L-2,4-diaminobutyrate (Ono *et al.*, 1999; Reshetnikov *et al.*, 2006), and its various interesting side reactions (Witt *et al.*, 2011; Moritz *et al.*, 2015), a full understanding of the enzyme activities of this ubiquitously distributed protein (Widderich *et al.*, 2014) is still lacking. Also, no structural information is currently available for proteins belonging to this family. To further the knowledge of ectoine synthase, we developed an overproduction and purification protocol for EctC based on a codon-optimized version of the *ectC* gene from *S. alaskensis* (*Sa*; Ting *et al.*, 2010) for its heterologous synthesis in *Escherichia coli* as a fusion protein modified with a *Strep*-tag II affinity tag at its carboxy-terminus. We demonstrate here that the recombinant *Sa*EctC-*Strep*-tag II protein can be efficiently overproduced and purified by affinity chromatography; it yields well diffracting crystals (to 1.0 Å resolution) that should allow the elucidation of the macromolecular structure of ectoine synthase.

2. Materials and methods

2.1. Macromolecule production

All recombinant DNA techniques followed routine procedures. The DNA sequence encoding the EctC protein (accession No. ABF54656) was retrieved from the genome

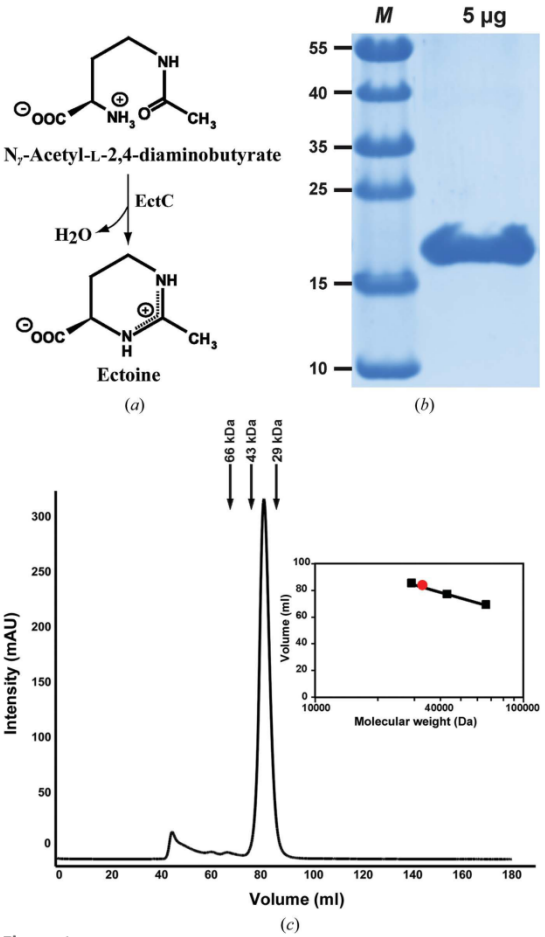


Figure 1
(*a*) Scheme of the EctC-catalyzed enzyme reaction. (*b*) SDS–PAGE analysis of the *Sa*EctC-*Strep*-tag II protein. Lane *M*, molecular-mass marker. (*c*) Gel-filtration analysis of the *Sa*EctC-*Strep*-tag II protein. A solution (3 mg ml^{-1}) of albumin (66 kDa), ovalbumin (43 kDa) and carboanhydrase (29 kDa) was used to standardize the HiLoad 16/600 Superdex 75 pg gel-filtration column. The inset shows the calculation of the molecular mass of the *Sa*EctC-*Strep*-tag II protein (closed red circle) with reference to the three marker proteins (closed black squares).

sequence of *S. alaskensis* strain RB2256 (Genome RefSeq No. NC_008048; Ting *et al.*, 2010). This template was used for the synthesis of a version of *ectC* that was codon-optimized for expression in *E. coli* (by Life Technologies, Darmstadt, Germany); the corresponding nucleotide sequence was deposited in the NCBI database under accession No. KR002036. The synthetic *ectC* DNA sequence present on the plasmid [pSynth_*ectC*(Sa)] provided by the supplier (Life Technologies) has been extended, at both its 5′- and 3′-ends, with a short DNA sequence containing LguI restriction sites to allow cloning into the pENTRY-IBA20 acceptor vector (IBA GmbH, Göttingen, Germany); this construction yielded plasmid pWN8. The DNA fragment carrying *ectC* was retrieved from pWN8 and inserted into the Esp3I site in the expression vector pASK-IBA3 (IBA GmbH) in such a way that the stop codon of *ectC* was removed and the 3′-end of the coding region was fused to a short DNA sequence encoding a nine-amino-acid (NWSHPQFEK) *Strep*-tag II affinity peptide. The resulting plasmid was named pWN12 and its nucleotide sequence was verified by DNA sequence analysis (carried out by Eurofins MWG, Ebersberg, Germany). Transcription of the *ectC* gene carried by pWN12 is driven by the *tet* promoter, the activity of which in turn is controlled by the TetR repressor, a genetic expression system that can be induced by adding anhydrotetracycline (AHT) to the growth medium.

2.1.1. Overexpression and purification of the *Sa*EctC-*Strep*-tag II protein. *E. coli* BL21 cells carrying the pWN12 expression plasmid were grown at 310 K in Minimal Medium A (MMA) with 0.5%(*w*/*v*) glucose as the carbon source, 0.5%(*w*/*v*) casamino acids, 1 m*M* $MgSO_4$ and 3 m*M* thiamine (Höppner *et al.*, 2014; Hoeppner *et al.*, 2014). The medium was supplemented with ampicillin (100 µg ml^{-1}) to select for the presence of plasmid pWN12 in the *E. coli* cells. The cells were grown in a 2 l Erlenmeyer flask (filled with 1 l medium) in an aerial shaker set to 180 rev min^{-1}. Overexpression of the plasmid-encoded *ectC* gene was initiated by adding the inducer AHT (final concentration of 0.2 mg per millilitre of culture) to the growth medium when the bacterial culture had reached an optical density (OD_{578}) of about 0.7. Subsequently, the growth temperature of the culture was reduced to 303 K, the speed of the aerial shaker was reduced to 100 rev min^{-1} and the cells were propagated for an additional 2 h to allow overproduction of the *Sa*EctC-*Strep*-tag II protein by the recombinant cells. Harvesting of the cells by centrifugation and their disruption by passage through a French Pressure Cell Press (SLM Aminco) followed the previously described procedures for the preparation of cell extracts containing the overproduced *Sa*EctD-*Strep*-tag II (Hoeppner *et al.*, 2014; Höppner *et al.*, 2014), except that the cells producing the *Sa*EctC-*Strep*-tag II protein were resuspended in a buffer consisting of 200 m*M* NaCl, 20 m*M* Tris pH 8 (buffer *A*). Cellular debris was removed by ultracentrifugation (60 min at 100 000*g* at 277 K) and the cleared supernatant was loaded onto a Strep-Tactin Superflow column (bed volume of 5 ml) that had been equilibrated with a solution of five bed volumes of buffer *A*. The column was then washed with ten column volumes of buffer *A*, and the *Sa*EctC-*Strep*-tag II protein was eluted from the affinity chromatography material with three column volumes of buffer *A* containing 2.5 m*M* desthiobiotin. Desthiobiotin and AHT were purchased from IBA GmbH, Göttingen, Germany. The concentration of the *Sa*EctC-*Strep*-tag II protein in the individual fractions eluted from the Strep-Tactin Superflow affinity column was measured with the Pierce BCA Protein Assay Kit (Thermo Scientific, Schwerte, Germany) and by using an extinction coefficient of 15 470 M^{-1} cm^{-1} at 280 nm and the molecular mass (16.3 kDa) of the full-length *Sa*EctC protein attached to the *Strep*-tag II affinity peptide. The purity and molecular mass of the *Sa*EctC-*Strep*-tag II protein was assessed by SDS–PAGE (15% polyacrylamide; Fig. 1*b*) using PageRuler Prestained Protein Ladder (Fischer Scientific GmbH, Schwerte, Germany) as a reference. The affinity-purified *Sa*EctC-*Strep*-tag II protein was routinely kept at 277 K. Prior to crystallization trials, the *Sa*EctC-*Strep*-tag II protein was concentrated to approximately 11 mg ml^{-1} (in 20 m*M* Tris, 200 m*M* NaCl pH 8) with Vivaspin 6 columns (Sartorius Stedim Biotech GmbH, Göttingen, Germany) with a 10 kDa molecular-weight cutoff. Macromolecule-production information is summarized in Table 1.

Table 1
Macromolecule-production information.

Source organism	*S. alaskensis* strain RB2256
DNA source	Synthetically made
Forward primer	GTGAAATGAATAGTTCGAC
Reverse primer	CGCAGTAGCGGTAAACGGC
Cloning vector	pSynth_*ectC*
Expression vector	pASK_IBA
Expression host	*E. coli*
Complete amino-acid sequence of the construct produced	MIVRNLGDIRKTDRNVRSDGWASARMLLKDDGMG-FSFHVTTLFAGSELRMHYQNHLEAVLVLKGTG-TIEDLATGEVHALRPGVMYALDDHDRHIVRPE-TDILTACVFNPPVTGREVHDESGAYPADPELA-REPVAADNWSHPQFEK

2.1.2. Size-exclusion chromatography of *Sa*EctC. Size-exclusion chromatography of the affinity-purified *Sa*EctC-*Strep*-tag II protein was performed on a HiLoad 16/600 Superdex 75 pg column (GE Healthcare, München, Germany) equilibrated and run in 200 m*M* NaCl, 20 m*M* Tris pH 8 under previously described conditions (Widderich *et al.*, 2014). 5 mg of the *Sa*EctC-*Strep*-tag II protein (in a volume of 1 ml) was loaded onto the size-exclusion column, which was standardized with albumin (66 kDa), ovalbumin (43 kDa) and carboanhydrase (29 kDa). These marker proteins were purchased from Sigma–Aldrich (Steinheim, Germany) and GE Healthcare (München, Germany).

2.2. Crystallization and preliminary X-ray analysis of *Sa*EctC

Several crystal hits for the *Sa*EctC-*Strep*-tag II protein were obtained using commercial screens from NeXtal (Qiagen, Hilden, Germany) and Molecular Dimensions (Suffolk, England) in Corning 3553 sitting-drop plates at 285 K. Homogeneous *Sa*EctC-*Strep*-tag II protein (0.1 µl of a solution at 11 mg ml^{-1}) was mixed with 0.1 µl reservoir solution and equilibrated against 50 µl reservoir solution. The most promising hit was found using a condition consisting of 0.05 *M*

research communications

calcium acetate, 0.1 *M* sodium acetate pH 4.5, 40%(*v*/*v*) 1,2-propanediol from the NeXtal Core IV suite (Qiagen, Hilden, Germany). This condition was optimized by grid screens around the initial condition and by the addition of several divalent cations. Large crystals were obtained either without any additive or on the addition of Ca^{2+} to the protein solution 30 min before the drops were placed. The crystals reached their maximum dimensions of about 50 × 50 × 70 µm after ten weeks. The crystals were harvested after overlaying the drop with 2 µl mineral oil and were flash-cooled in liquid nitrogen. Crystallization information is summarized in Table 2.

2.3. Data collection and processing

Data sets were collected from a single crystal of the *Sa*EctC-*Strep*-tag II protein on beamline ID29 at the ERSF, Grenoble, France at 100 K using a Pilatus detector. After the initial diffraction tests, the data-collection strategy was first calculated using the *EDNA* software available at the beamline (Incardona *et al.*, 2009) and the subsequently collected data sets were processed using the *XDS* package (Kabsch, 2010*a*) and scaled with *XSCALE* (Kabsch, 2010*b*). Data-collection and processing statistics are summarized in Table 3.

3. Results and discussion

To provide the large amounts of ectoine synthase (EctC) required for structural analysis, we obtained a codon-optimized version of the *ectC* gene from the cold-adapted marine bacterium *S. alaskensis* (*Sa*; Ting *et al.*, 2010) for expression in *E. coli*. We chose the *Sa*EctC protein for our studies since we recently successfully determined the crystal structures of the apo and liganded forms of ectoine hydroxylase (EctD; Höppner *et al.*, 2014; Hoeppner *et al.*, 2014) from this microorganism. Overproduction of the *Sa*EctC-*Strep*-tag II protein in *E. coli* strain BL21 (pWN12) typically yielded approximately 40–50 mg of protein per litre of bacterial cell culture. As evaluated by SDS–PAGE, the *Sa*EctC-*Strep*-tag II protein was at least 98% pure as judged by optical inspection of a 15% SDS–polyacrylamide gel (Fig. 1*b*). Its electrophoretic mobility on the SDS–polyacrylamide gel was consistent with the theoretically calculated molecular mass (16.3 kDa, including the attached *Strep*-tag II peptide) of the recombinant protein (Fig. 1*b*).

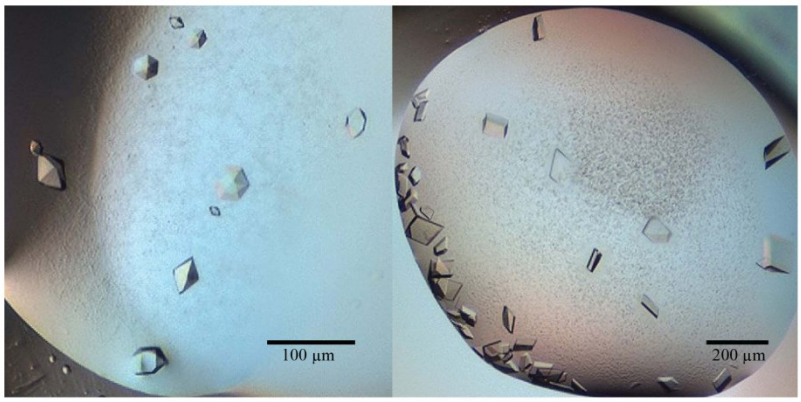
Table 2
Crystallization.

Method	Sitting-drop vapour diffusion
Plate type	Corning 3553
Temperature (K)	285
Protein concentration (mg ml^{-1})	11
Buffer composition of protein solution	200 m*M* NaCl, 20 m*M* Tris pH 8 plus 5 m*M* Ca^{2+} prior to crystallization
Composition of reservoir solution	0.05 *M* calcium acetate, 0.1 *M* sodium acetate pH 4.5, 40%(*v*/*v*) 1,2-propanediol
Volume and ratio of drop	1:1
Volume of reservoir (µl)	300

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	ID29, ESRF, Grenoble, France
Wavelength (Å)	0.97625
Temperature (K)	100
Detector	Pilatus-6M
Crystal-to-detector distance (mm)	164.35
Rotation range per image (°)	0.05
Total rotation range (°)	360
Exposure time per image (s)	0.02
Space group	$P3_221$
a, *b*, *c* (Å)	72.715, 72.715, 52.330
α, β, γ (°)	90.0, 90.0, 120.0
Mosaicity (°)	0.06
Resolution range (Å)	30–1.0
Total No. of reflections	1364845 (46924)
No. of unique reflections	882245 (7645)
Completeness (%)	98.0 (83.7)
Multiplicity	16.5 (6.1)
$\langle I/\sigma(I)\rangle$	16.2 (2.0)
$R_{r.i.m.}$ (%)	8.9 (115.4)
$CC_{1/2}$	99.9 (43.7)
Overall *B* factor from Wilson plot (Å^2)	10.2



(*a*) (*b*)

Figure 2
Crystals of *Sa*EctC. (*a*) Initial and (*b*) optimized crystals

The homogeneity of the affinity-purified *Sa*EctC-*Strep*-tag II protein was assessed analytically by size-exclusion chromatography. The vast majority of the protein eluted from a HiLoad 16/600 Superdex 75 pg column in a single symmetrical peak, and only a very minor portion of the *Sa*EctC-*Strep*-tag II preparation was present as aggregates (Fig. 1*c*). It eluted from the size-exclusion chromatography column as a species with a calculated molecular mass of 32.7 kDa, indicating that the *Sa*EctC-*Strep*-tag II protein forms a dimer in solution (Fig. 1*c*). This finding is in line with the purified native EctC protein from *H. elongata*, which is probably also a dimer in solution (Ono *et al.*, 1999). Taken together, we conclude from these experiments that the recombinant *Sa*EctC-*Strep*-tag II protein can be effectively overproduced in *E. coli* and purified with good yields and that these protein preparations are by and large homogeneous.

Well diffracting crystals of the *Sa*EctC-*Strep*-tag II protein were obtained with 0.06 *M* calcium acetate, 0.08 *M* sodium acetate pH 4.5, 38–42%(*v*/*v*) 1,2-propanediol using the sitting-drop vapour-diffusion method (Fig. 2). Native data sets for the *Sa*EctC-*Strep*-tag II crystal species were collected at 100 K to a maximum resolution of 1.0 Å (Fig. 3). Preliminary data processing using the *XDS* package displayed $P3_221$ symmetry (Table 1). The V_M value was calculated to be 2.32 Å^3 Da^{-1} with a monomer in the asymmetric unit, with a solvent content of 47% (Matthews, 1968).

Members of the cupin superfamily of proteins possess a rather well conserved common β-barrel core structural scaffold (Uberto & Moomaw, 2013), but our attempts to solve the *Sa*EctC crystal structure by molecular replacement using the crystal structure of the Zn^{2+}-containing cupin RemF from *Streptomyces resistomycificus* (PDB entry 3ht1; Silvennoinen *et al.*, 2009), a polyketide cyclase involved in the biosynthesis of the metabolite resistomycin, were unsuccessful. Notably, the *Sa*EctC protein possesses a single Cys residue that should be ideally suited for heavy-atom binding to obtain crystals of the *Sa*EctC-*Strep*-tag II protein that can be used to obtain initial phases *via* MAD experiments in order to solve the crystal structure of ectoine synthase.

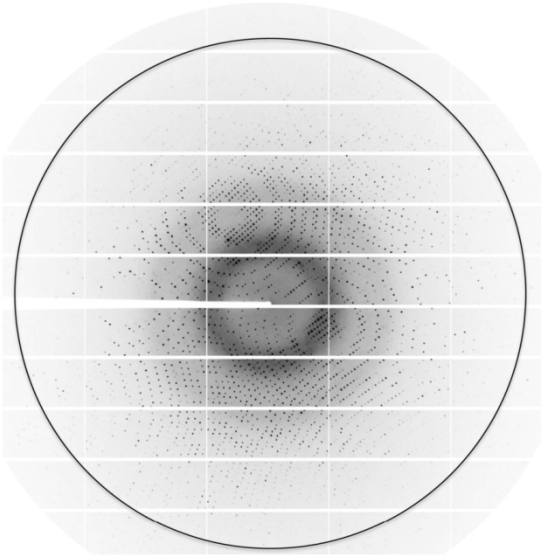


Figure 3
Diffraction images of *Sa*EctC (oscillation width 0.5°). The images were used to calculate the data-collection strategy using the *EDNA* software (Incardona *et al.*, 2009). The circle represents a resolution of 1.2 Å

Acknowledgements

We acknowledge the European Synchrotron Radiation Facility for the provision of synchrotron-radiation resources, especially the staff of the ID29 beamline at the ESRF, Grenoble, France. We gratefully acknowledge the financial support from the Heinrich Heine University, the Institute for Biochemistry at the University of Düsseldorf, the DFG-funded SFB-987 (University of Marburg) and the LOEWE Center for Synthetic Microbiology (Marburg) for this study. We thank Lutz Schmitt for his interest in and support of this collaborative project. NW is the recipient of a PhD fellowship from the International Max Planck Research School for Environmental, Cellular and Molecular Microbiology (IMPRS-Mic Marburg). We greatly appreciate the kind help of Vickie Koogle in the language editing of our manuscript and thank Jochen Sohn for his expert technical assistance.

References

Bursy, J., Kuhlmann, A. U., Pittelkow, M., Hartmann, H., Jebbar, M., Pierik, A. J. & Bremer, E. (2008). *Appl. Environ. Microbiol.* **74**, 7286–7296.

Galinski, E. A., Pfeiffer, H. P. & Trüper, H. G. (1985). *Eur. J. Biochem.* **149**, 135–139.

García-Estepa, R., Argandoña, M., Reina-Bueno, M., Capote, N., Iglesias-Guerra, F., Nieto, J. J. & Vargas, C. (2006). *J. Bacteriol.* **188**, 3774–3784.

Graf, R., Anzali, S., Buenger, J., Pfluecker, F. & Driller, H. (2008). *Clin. Dermatol.* **26**, 326–333.

Hoeppner, A., Widderich, N., Bremer, E. & Smits, S. H. J. (2014). *Acta Cryst.* **F70**, 493–496.

Höppner, A., Widderich, N., Lenders, M., Bremer, E. & Smits, S. H. J. (2014). *J. Biol. Chem.* **289**, 29570–29583.

Incardona, M.-F., Bourenkov, G. P., Levik, K., Pieritz, R. A., Popov, A. N. & Svensson, O. (2009). *J. Synchrotron Rad.* **16**, 872–879.

Kabsch, W. (2010*a*). *Acta Cryst.* D**66**, 125–132.

Kabsch, W. (2010*b*). *Acta Cryst.* D**66**, 133–144.

Kempf, B. & Bremer, E. (1998). *Arch. Microbiol.* **170**, 319–330.

Kuhlmann, A. U., Hoffmann, T., Bursy, J., Jebbar, M. & Bremer, E. (2011). *J. Bacteriol.* **193**, 4699–4708.

Lippert, K. & Galinski, E. A. (1992). *Appl. Microbiol. Biotechnol.* **37**, 61–65.

Louis, P. & Galinski, E. A. (1997). *Microbiology*, **143**, 1141–1149.

Manzanera, M., García de Castro, A., Tøndervik, A., Rayner-Brandes, M., Strøm, A. R. & Tunnacliffe, A. (2002). *Appl. Environ. Microbiol.* **68**, 4328–4333.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Moritz, K. D., Amendt, B., Witt, E. M. H. J. & Galinski, E. A. (2015). *Extremophiles*, **19**, 87–99.

research communications

Ono, H., Sawada, K., Khunajakr, N., Tao, T., Yamamoto, M., Hiramoto, M., Shinmyo, A., Takano, M. & Murooka, Y. (1999). *J. Bacteriol.* **181**, 91–99.
Pastor, J. M., Salvador, M., Argandoña, M., Bernal, V., Reina-Bueno, M., Csonka, L. N., Iborra, J. L., Vargas, C., Nieto, J. J. & Cánovas, M. (2010). *Biotechnol. Adv.* **28**, 782–801.
Reshetnikov, A. S., Khmelenina, V. N. & Trotsenko, Y. A. (2006). *Arch. Microbiol.* **184**, 286–297.
Silvennoinen, L., Sandalova, T. & Schneider, G. (2009). *FEBS Lett.* **583**, 2917–2921.
Tanne, C., Golovina, E. A., Hoekstra, F. A., Meffert, A. & Galinski, E. A. (2014). *Front. Microbiol.* **5**, 150.
Ting, L., Williams, T. J., Cowley, M. J., Lauro, F. M., Guilhaus, M., Raftery, M. J. & Cavicchioli, R. (2010). *Environ. Microbiol.* **12**, 2658–2676.
Uberto, R. & Moomaw, E. W. (2013). *PLoS One*, **8**, e74477.
Widderich, N., Höppner, A., Pittelkow, M., Heider, J., Smits, S. H. J. & Bremer, E. (2014). *PLoS One*, **9**, e93809.
Witt, E. M. H. J., Davies, N. W. & Galinski, E. A. (2011). *Appl. Microbiol. Biotechnol.* **91**, 113–122.

VI. Publication 6

The sixth publication is entitled ‘Strangers in the archaeal world: osmostress-responsive biosynthesis of ectoine and hydroxyectoine by the marine thaumarchaeon *Nitrosopumilus maritimus*’. It was published in ‘Environmental Microbiology’ (doi: 10.1111/1462-2920.13156).

Widderich, N., Czech, L., Elling, F.J., Könneke, M., Stöveken, N., Pittelkow, M., Riclea, R., Dickschat, J.S., Heider, J., and Bremer, E. Strangers in the archaeal world: osmostress-responsive biosynthesis of ectoine and hydroxyectoine by the marine thaumarchaeon *Nitrosopumilus maritimus*. *Environ Microbiol*. (2016). 18: 1227–1248.

In this paper an updated phylogenetic distribution of ectoine biosynthetic genes was assessed and revealed the presence of ectoine gene clusters in several members of the *Archaea*. Genetic and molecular approaches in the archaeal representative *N. maritimus* demonstrated that its ectoine gene cluster is encoded in an operon whose expression is responsive to osmotic stress, as is the intracellular accumulation of ectoines. Determination of the salient biochemical and kinetic properties of the ectoine synthase and the ectoine hydroxylase from this archaeon illustrate that their characteristics resemble those of their bacterial counterparts. In addition, a gene encoding a mechanosensitive channel has been identified to be co-transcribed with the *ect*-operon and its functionality has been proven through genetic complementation assays. This study thus provides an updated phylogenetic affiliation and illustrates that ectoine biosynthesis in *Archaea* occurs similar to that in *Bacteria*. However, ectoine/hydroxyectoine production is rare in *Archaea* and the presented results suggest that the corresponding genes were acquired by lateral gene transfer.

Personal contribution:
I performed the database searches and assessed the phylogenetic distribution of the ectoine biosynthetic genes. I conducted production and purification of the EctC and EctD enzymes and determined their salient biochemical and kinetic properties by enzyme assays. I mapped the transcriptional organization of the ectoine gene cluster, assessed its osmo-responsive expression by qRT-PCR and analyzed the intracellular ectoine content. I analyzed the data (with the other authors) and prepared Figures and Tables (together with Laura Czech). I also contributed to the writing of the manuscript together with my PhD supervisor Erhard Bremer.

environmental microbiology

Environmental Microbiology (2016) doi:10.1111/1462-2920.13156

Strangers in the archaeal world: osmostress-responsive biosynthesis of ectoine and hydroxyectoine by the marine thaumarchaeon *Nitrosopumilus maritimus*

Nils Widderich,[1†] Laura Czech,[1†] Felix J. Elling,[2] Martin Könneke,[2] Nadine Stöveken,[1,3] Marco Pittelkow,[1] Ramona Riclea,[4,5] Jeroen S. Dickschat,[4,5] Johann Heider[1,3] and Erhard Bremer[1,3*]

[1]*Laboratory for Molecular Microbiology, Department of Biology, Philipps-University, Karl-von-Frisch Str. 8, D-35043 Marburg, Germany.*

[2]*Organic Geochemistry Group, MARUM – Center for Marine Environmental Sciences, University of Bremen, PO Box 330 440, D-28334 Bremen, Germany.*

[3]*LOEWE-Center for Synthetic Microbiology, Philipps-University Marburg, Hans-Meerwein-Str. 6, D-35043 Marburg, Germany.*

[4]*Kekulé-Institut for Organic Chemistry and Biochemistry, Friedrich-Wilhelms-University Bonn, Gerhard-Domagk Str. 1, D-53121 Bonn, Germany.*

[5]*Institute of Organic Chemistry, TU Braunschweig, Hagenring 30, D-38106 Braunschweig, Germany.*

Summary

Ectoine and hydroxyectoine are compatible solutes widely synthesized by members of the *Bacteria* to cope with high osmolarity surroundings. Inspection of 557 archaeal genomes revealed that only 12 strains affiliated with the *Nitrosopumilus*, *Methanothrix* or *Methanobacterium* genera harbour ectoine/hydroxyectoine gene clusters. Phylogenetic considerations suggest that these *Archaea* have acquired these genes through horizontal gene transfer events. Using the Thaumarchaeon '*Candidatus* Nitrosopumilus maritimus' as an example, we demonstrate that the transcription of its *ectABCD* genes is osmotically induced and functional since it leads to the production of both ectoine and hydroxyectoine. The ectoine synthase and the ectoine hydroxylase were biochemically characterized, and their properties resemble those of their counterparts from *Bacteria*. Transcriptional analysis of osmotically stressed '*Ca.* N. maritimus' cells demonstrated that they possess an ectoine/hydroxyectoine gene cluster (*hyp-ectABCD-mscS*) different from those recognized previously since it contains a gene for an MscS-type mechanosensitive channel. Complementation experiments with an *Escherichia coli* mutant lacking all known mechanosensitive channel proteins demonstrated that the (*Nm*)MscS protein is functional. Hence, '*Ca.* N. maritimus' cells cope with high salinity not only through enhanced synthesis of osmostress-protective ectoines but they already prepare themselves simultaneously for an eventually occurring osmotic down-shock by enhancing the production of a safety-valve (*Nm*MscS).

Received 31 July, 2015; revised 19 November, 2015; accepted 27 November, 2015. *For correspondence. E-mail bremer@staff.uni-marburg.de; Tel: (+49) 6421 2821529; Fax: (+49) 6421 2828979.
†These authors contributed equally.

Introduction

Microorganisms are masters of change. Their genetically encoded and physiologically mediated adaptive responses allow them to cope with nutrient limitations and a multitude of cellular and environmentally imposed stresses that otherwise would impair growth or threaten their viability (Storz and Hengge-Aronis, 2000). Increases in the external salinity or osmolarity are such types of stress and are encountered by essentially all free-living microorganisms (Csonka, 1989; Bremer and Krämer, 2000; Roesser and Müller, 2001). They trigger the rapid and passive outflow of water from the cell, thereby causing a drop in vital turgor and an increase in molecular crowding of the cytoplasm that in extreme cases causes growth arrest or even death (Record *et al.*, 1998; Wood, 2011). In the course of evolution, microorganisms have developed two different, but not necessarily mutually exclusive (Deole *et al.*, 2013; Oren, 2013; Becker *et al.*, 2014; Youssef *et al.*, 2014), mechanisms to counteract high osmolarity-instigated water efflux, the so-called *salt in* and the *salt out* strategies (Galinski and Trüper, 1994; Kempf and Bremer, 1998). Both are based on an active raising by the bacterial cell of the osmotic potential of its cytoplasm, thereby indirectly creating an osmotic driving

force to promote water retention and re-entry (Csonka, 1989; Record *et al.*, 1998; Bremer and Krämer, 2000; Roesser and Müller, 2001; Wood, 2011).

Microorganisms using the *salt in* strategy accumulate molar concentrations of K^+ and Cl^- ions through transport processes to balance the osmotic gradient across their cytoplasmic membrane (Galinski and Trüper, 1994; Oren, 2013). This is considered as an energetically favourable strategy for microorganisms to adapt to persistent high-salinity environments (Oren, 2011). However, it comes with a penalty since the entire proteome had to be adjusted in the course of evolution to sustained high ionic strength in order to keep proteins soluble and functional (Coquelle *et al.*, 2010; Deole *et al.*, 2013; Talon *et al.*, 2014). As a consequence, only a limited number of microbial species use the *salt in* strategy since it also often restricts the types of habitats that these microorganisms can populate (Galinski and Trüper, 1994; Oren, 2013).

Microorganisms that use the *salt out* strategy amass, either through synthesis or uptake, a restricted set of organic osmolytes, the compatible solutes (Csonka, 1989; da Costa *et al.*, 1998; Kempf and Bremer, 1998; Roesser and Müller, 2001; Wood *et al.*, 2001). This strategy affords a flexible physiological response to both sustained high-salinity surroundings and to osmotic changes in those ecosystems where the salinity fluctuates more often (Galinski and Trüper, 1994; Bremer and Krämer, 2000). Compatible solutes are operationally defined as small, highly water-soluble organic osmolytes that are fully compliant with cellular biochemistry and physiology (Brown, 1976; Bolen and Baskakov, 2001; Wood, 2011). Their benign nature allows their accumulation to exceedingly high intracellular levels without disturbing cellular physiology and biochemistry. The accumulation of compatible solutes occurs in a fashion that is correlated with the degree of the osmotic stress imposed onto the microbial cell (Kuhlmann and Bremer, 2002; Saum and Müller, 2008; Brill *et al.*, 2011; Hoffmann *et al.*, 2013). Compatible solutes used by members of the *Bacteria* are typically uncharged or zwitterionic (Csonka, 1989; da Costa *et al.*, 1998; Kempf and Bremer, 1998; Klähn and Hagemann, 2011). In contrast, members of the *Archaea* often produce charged derivatives of these types of osmolytes and also synthesize types of compatible solutes that are normally not found in *Bacteria* (da Costa *et al.*, 1998; Roesser and Müller, 2001; Roberts, 2004; Müller *et al.*, 2005; Empadinhas and da Costa, 2011).

Here, we focus on the tetrahydropyrimidine ectoine and its derivative 5-hydroxyectoine (Galinski *et al.*, 1985; Inbar and Lapidot, 1988) (Fig. 1A), widely used compatible solutes in the microbial world (Pastor *et al.*, 2010; Widderich *et al.*, 2014a). High-osmolarity environments trigger enhanced expression of the ectoine (*ectABC*) and hydroxyectoine (*ectD*) biosynthetic genes (Louis and Galinski, 1997; Prabhu *et al.*, 2004; Garcia-Estepa *et al.*, 2006; Bursy *et al.*, 2007), thereby leading to high-level synthesis of these compounds under osmotically unfavourable conditions (Kuhlmann and Bremer, 2002; Calderon *et al.*, 2004; Bursy *et al.*, 2007; Salvador *et al.*, 2015). Genetic disruption of the ectoine biosynthetic genes causes an osmotic-sensitive growth phenotype, underscoring the physiological role of ectoines as effective microbial osmostress protectants (Canovas *et al.*, 1997; Kol *et al.*, 2010).

Synthesis of ectoine progresses from the precursor L-aspartate-β-semialdehyde, a central intermediate in microbial amino acid metabolism and cell wall synthesis (Lo *et al.*, 2009; Stöveken *et al.*, 2011). It comprises the sequential activities of three enzymes: L-2,4-diaminobutyrate transaminase (EctB; EC 2.6.1.76), L-2,4-diaminobutyrate acetyltransferase (EctA; EC 2.3.1.178) and ectoine synthase (EctC; EC 4.2.1.108) (Louis and Galinski, 1997; Ono *et al.*, 1999b) (Fig. 1A). A considerable number of ectoine producers also synthesize an ectoine derivative, 5-hydroxyectoine, through a stereospecific reaction that is catalysed by the ectoine hydroxylase (EctD) (EC 1.14.11) (Bursy *et al.*, 2007; Höppner *et al.*, 2014; Widderich *et al.*, 2014a,b) (Fig. 1A). Hydroxyectoine exhibits stress-protective properties that are partially different from and sometimes exceed those of ectoine, in particular with respect to temperature and desiccation stress (Manzanera *et al.*, 2002; Garcia-Estepa *et al.*, 2006; Bursy *et al.*, 2008; Tanne *et al.*, 2014).

A recent survey of bacterial and archaeal genome sequences revealed that ectoine and hydroxyectoine biosynthetic genes occur widely in *Bacteria*, are rarely found in *Archaea*, and do not exist in *Eukarya* (Widderich *et al.*, 2014a). In a few isolated incidents, the presence of ectoine/hydroxyectoine biosynthetic genes in *Archaea* has been noted, e.g. in the context of the annotation of the genome sequences of the Thaumarchaeon '*Candidatus* Nitrosopumilus maritimus' SCM1 (Walker *et al.*, 2010) and of the methanogen *Methanobacterium formicicum* BRM9 (Kelly *et al.*, 2014). However, it is unclear whether these genes are functional, whether the properties of the key enzymes for ectoine/hydroxyectoine biosynthesis, the ectoine synthase (EctC) and the ectoine hydroxylase (EctD) (Fig. 1A) are different from those of their bacterial counterparts, and whether the transcription of the archaeal *ect* genes is increased in response to osmotic stress as is observed in members of the *Bacteria* (Kuhlmann and Bremer, 2002; Calderon *et al.*, 2004; Bursy *et al.*, 2007; Kuhlmann *et al.*, 2008; Salvador *et al.*, 2015).

Here, we address these questions by interrogating the rapidly expanding number of archaeal genome sequences through bioinformatics to provide a comprehensive overview on the occurrence of ectoine/

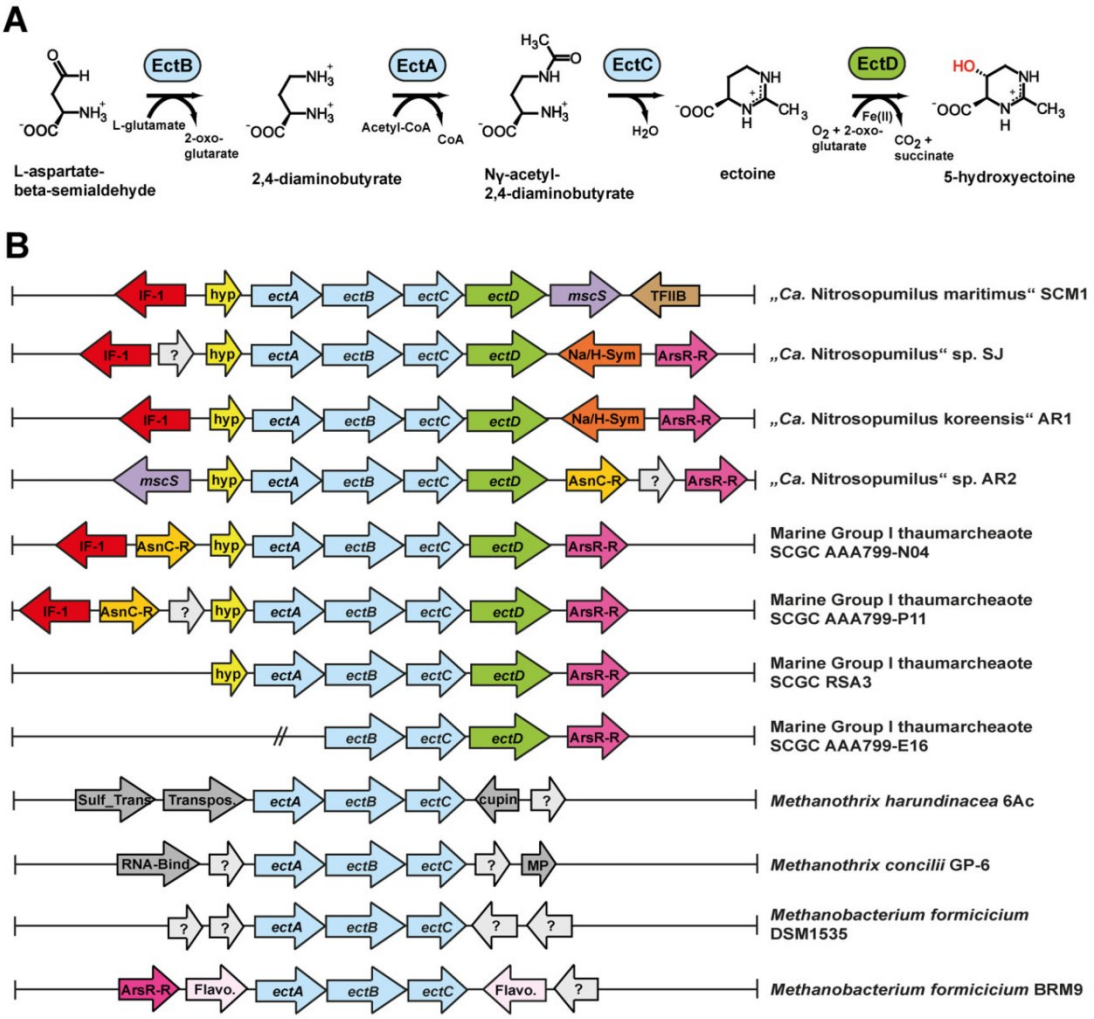


Fig. 1. (A) Biochemical steps required for the biosynthesis of ectoine and 5-hydroxyectoine and (B) genetic organization of *ectABC(D)* gene cluster and their flanking regions in archaeal genomes. In the marine group I thaumarchaeote SCGC AAA799-E16, no *ectA* gene was present; however, this genome sequence was assembled by a metagenomic approach making it likely that this thaumarchaeote also possesses a fully *ectABCD* gene cluster.

hydroxyectoine gene clusters in this domain of life. We used the marine Thaumarchaeon '*Ca.* Nitrosopumilus maritimus' SCM1 (Könneke *et al.*, 2005; Walker *et al.*, 2010), the first cultured representative of the globally abundant ammonia-oxidizing *Archaea* (AOA) (Stahl and de la Torre, 2012; Offre *et al.*, 2013; Bayer *et al.*, 2015; Gubry-Rangin *et al.*, 2015), as a model system to study ectoine/hydroxyectoine production in detail. We demonstrate for the first time in any archaeon that its *ectABCD* gene cluster is functionally expressed in response to high salinity, and that the biochemical properties of key enzymes (EctC, EctD) for the synthesis of ectoine and hydroxyectoine resemble those of their bacterial counterparts. Most interestingly, we found that the *ectABCD* genes of '*Ca.* N. maritimus' SCM1 are not only osmotically inducible but that they are also co-transcribed with a gene that encodes a functional mechanosensitive channel of the MscS family. These safety valves are ubiquitously used by both *Bacteria* and *Archaea* to withstand rapid osmotic downshifts (Kloda and Martinac, 2002; Booth and Blount, 2012; Wilson *et al.*, 2013; Booth, 2014). Hence, the osmotically regulated *ectABCD-mscS* transcriptional

unit is a sophisticated genetic device that allows high-salinity challenged '*Ca.* N. maritimus' SCM1 cells to sequentially cope with increases and decreases in the external osmolarity of their marine and estuarine habitat.

Results

Assessing the distribution of ectoine/hydroxyectoine biosynthetic genes in Bacteria and Archaea

To assess the occurrence and taxonomic distribution of ectoine biosynthetic genes, we used the database of the Joint Genome Institute (JGI) of the US Department of Energy (http://jgi.doe.gov/) (Nordberg *et al.*, 2013) and the EctC protein from '*Ca.* Nitrosopumilus maritimus' SCM1 (Walker *et al.*, 2010) as the query sequence. While the EctA (2,4-diaminobutyrate acetyltransferase) and EctB (L-2,4-diaminobutyrate transaminase) enzymes have counterparts in microbial biosynthetic pathways not related to ectoine biosynthesis, EctC can be regarded as a diagnostic enzyme for ectoine producers (Widderich *et al.*, 2014a). Of note is that a restricted number of microorganisms exists that possesses *ectC*-type genes but lack the corresponding *ectAB* genes. The functional relevance of these orphan EctC-type proteins for ectoine biosynthesis is currently not fully understood (Kurz *et al.*, 2010; Widderich *et al.*, 2014a).

At the time of the BLAST search (08.07.2015), 27.232 completed and partially completed genome sequences of members of the *Bacteria* were represented in the JGI microbial database and 557 genome sequences of *Archaea* had been deposited. Among these 27.789 microbial genome sequences, 1297 hits to EctC-type proteins were found. After removing redundant entries of closely related strains (e.g. there are 181 genomes of strain of *Vibrio cholerae* represented that each possesses an *ect* gene cluster), the curated data set comprised 723 EctC-type proteins; 711 originate from *Bacteria* and only 12 were derived from *Archaea*. We note in this context that a considerable number of EctC-type proteins are misannotated in the database as L-mannose-6-phosphate isomerases (despite the typical localization of the corresponding genes within *ectABC(D)* gene clusters). Often, they are also referred to as RmlC-type cupins, an error that might originate from the bioinformatics assignment of the EctC protein to the RmlC subgroup of the cupin superfamily (Dunwell *et al.*, 2001). The RmlC enzyme participates in the L-rhamnose biosynthetic pathway, serves as a carbohydrate epimerase (EC 5.1.3.13) (Dong *et al.*, 2007), and therefore catalyses an enzymatic reaction quite different from that of the ectoine synthase EctC (EC 4.2.1.108) (Ono *et al.*, 1999b) (Fig. 1A).

We then focused our further analysis on the genome sequences of those 12 *Archaea* that were identified by our database analysis to harbour *ectC*-type genes (Fig. 1B). In each of the corresponding genomes, the *ectC* gene is part of either an *ectABC* operon, or is embedded in an *ect* gene cluster that also comprises *ectD*. Eight members affiliated with the genus *Nitrosopumilus* or with Marine Group I.1a Thaumarchaeota (Pester *et al.*, 2011) were represented among the 12 *ect*-containing *Archaea*, and in each of these gene clusters, a copy of the ectoine hydroxylase gene (*ectD*) was also found (Fig. 1B). Two species of the genus *Methanothrix*, *Methanothrix harundinacea* 6Ac and *Methanothrix concilii* GP6 were represented, and two strains of *Methanobacterium* were found as well. In the genome sequence of these four methanogens, a full *ectABC* gene cluster was present, but each of them lacked the ectoine hydroxylase *ectD* gene (Fig. 1B), and such a gene was also not present anywhere else in their genome. The ectoine hydroxylase (EctD) is an oxygen-dependent enzyme (Bursy *et al.*, 2007; Höppner *et al.*, 2014; Widderich *et al.*, 2014b). Since members of the *Methanothrix* and *Methanobacterium* genera are all strict anaerobes, the absence of *ectD* in the genomes of these methanogens is readily understandable. On the other hand, the presence of *ectD* genes in '*Ca.* N. maritimus' SCM1, other members of the *Nitrosopumilus* genus and related Marine Group I Thaumarchaeota (Fig. 1B) is consistent with the lifestyle of these *Archaea* as oxygen-dependent nitrifying microorganisms (Könneke *et al.*, 2005; Pester *et al.*, 2011; Stieglmeier *et al.*, 2014).

'Ca. N. maritimus' SCM1 synthesizes both ectoine and hydroxyectoine in response to osmotic stress

The presence of a particular gene cluster in a microbial genome does not necessarily imply that it is also functional. To study whether any of the detected archaeal *ect* gene clusters were functionally expressed, we chose the Thaumarchaeon '*Ca.* N. maritimus' SCM1 as a model system (Könneke *et al.*, 2005; Walker *et al.*, 2010). '*Candidatus* Nitrosopumilus maritimus' SCM1 was originally isolated from a tropical fish tank of an aquarium in Seattle (USA) and represents the first cultured ammonium oxidizer within the domain *Archaea* (Könneke *et al.*, 2005). '*Ca.* N. maritimus' SCM1 belongs to the phylum Thaumarchaeota (Brochier-Armanet *et al.*, 2008; Stieglmeier *et al.*, 2014), which are among the most abundant microorganisms on Earth. They are ubiquitous in the Ocean (Stahl and de la Torre, 2012), found in estuarine sediments (Zhang *et al.*, 2015) and are also ubiquitous in terrestrial habitats (Gubry-Rangin *et al.*, 2015). Being adapted to extreme oligotrophic conditions (Martens-Habbena *et al.*, 2009; Könneke *et al.*, 2014), ammonia-oxidizing Thaumarchaeota are the predominant nitrifyers in the ocean and contribute significantly to the marine nitrogen cycle (Offre *et al.*, 2013). The predicted EctABCD proteins of '*Ca.* N. maritimus' SCM1

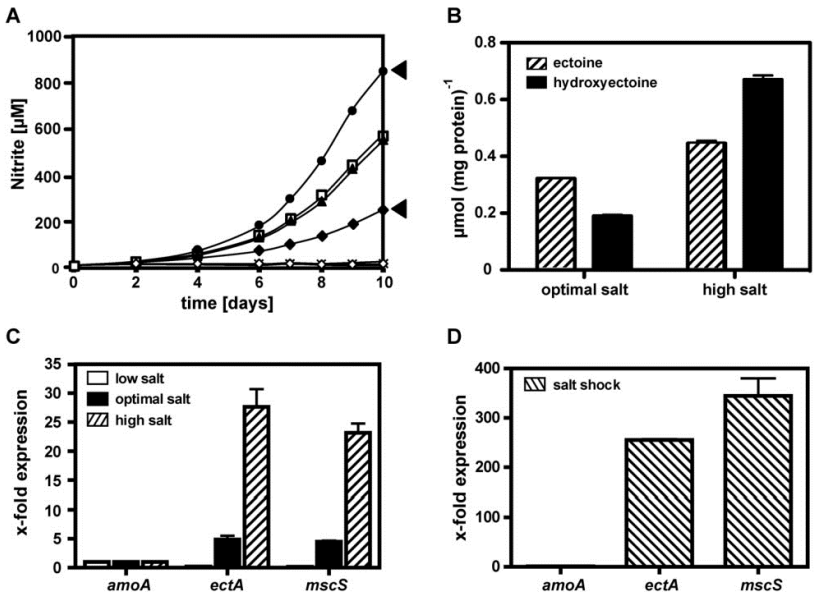


Fig. 2. Growth of '*Ca.* N. maritimus' SCM1 in the presence of different NaCl concentrations, its production of ectoines and the expression of its *ect* operon in response to different NaCl concentrations.
A. Growth of '*Ca.* N. maritimus' SCM1 cells in the presence of 0 M (closed diamonds), 0.15 M (cross), 0.32 M (open triangles), 0.48 M (closed triangles), 0.63 M (closed circles), 0.79 M (open squares), 0.94 M (closed diamonds) and 1.27 M (open diamonds) NaCl was monitored by nitrite production. Optimal sodium chloride concentration (0.63 M; closed circles) and high salt concentration (0.94 M; closed diamonds) in the growth medium of those cells used for the analysis of their ectoine and hydroxyectoine content are indicated by black arrowheads.
B. Intracellular ectoine and hydroxyectoine content per total amount of protein as assessed by HPLC measurements of '*Ca.* N. maritimus' SCM1 cells grown under optimal and high salt concentrations. Changes in expression of the '*Ca.* N. maritimus' SCM1 *ectA* and *mscS* genes under (C) low, optimal and high salt concentrations and (D) after an osmotic shock as assessed by qRT-PCR are depicted. Relative expression was determined using the *amoA* transcript as a reference.

exhibited a degree of amino acid sequence identity of 41%, 53%, 52% and 47% in comparison with the functionally studied ectoine and hydroxyectoine biosynthetic enzymes of *Halomonas elongata* (Ono *et al.*, 1999b; Widderich *et al.*, 2014a), the natural microbial cell factory for the industrial-scale production of ectoines (Schwibbert *et al.*, 2011; Kunte *et al.*, 2014).

'*Ca.* N. maritimus' SCM1 can grow over a wide range of salinities (Elling *et al.*, 2015). To assess whether it would synthesize both ectoine and hydroxyectoine in response to salt stress, '*Ca.* N. maritimus' SCM1 was grown in a chemically defined medium containing 26 g l^{-1} NaCl (0.63 M NaCl); this medium has a measured osmolarity of 928 mOsm. To impose additional osmotic stress on '*Ca.* N. maritimus' SCM1 cells, we also grew them in the same medium but with increased sodium chloride (NaCl) concentration (48 g l^{-1}) (0.94 M NaCl); this medium had a measured osmolarity of 1.598 mOsm. The increased salinity of the medium had a significant negative impact on the growth of '*Ca.* N. maritimus' SCM1 (Fig. 2A). Under both growth conditions, ectoine and hydroxyectoine were detected by high-performance liquid chromatography (HPLC) analysis, demonstrating that the *ectABCD* gene cluster of '*Ca.* N. maritimus' SCM1 was functional. Since '*Ca.* N. maritimus' SCM 1 is a marine microorganism (Könneke *et al.*, 2005), optimal growth conditions occur only at a substantial osmolarity (26 g NaCl l^{-1}) (0.63 M NaCl). Hence, it was not surprising that we already found ectoine and hydroxyectoine in these cells; their total ectoine and hydroxyectoine content were 0.32 µmol mg protein^{-1} and 0.19 µmol mg protein^{-1} respectively. These values increased to 0.44 µmol mg protein^{-1} of ectoine and 0.67 µmol mg protein^{-1} of hydroxyectoine when the cells were grown under heightened osmotic stress in a medium with a sodium chloride concentration to 48 g l^{-1} (0.94 M NaCl). Consequently, the level of synthesis of these compatible solutes by '*Ca.* N. maritimus' SCM1 is responsive to increased osmotic stress (Fig. 2B).

Biochemical properties of the ectoine synthase and ectoine hydroxylase from 'Ca. N. maritimus' SCM1

EctC and EctD are the key enzymes for ectoine and hydroxyectoine synthesis respectively (Fig. 1A) (Ono *et al.*, 1999b; Bursy *et al.*, 2007; Widderich *et al.*, 2014a). We therefore wondered whether the biochemical properties of the EctC and EctD enzymes derived from *Archaea* possess characteristics similar to their bacterial counterparts, or are different. Consequently, we assessed the properties and kinetics of the EctC and EctD enzymes from the ectoine/hydroxyectoine producer '*Ca.* N. maritimus' SCM1. To provide the substantial amounts of proteins required for biochemical studies, we obtained synthetic, codon-optimized versions of the *ectC* and *ectD* genes from '*Ca.* N. maritimus' SCM1 (Walker *et al.*, 2010) and inserted them into the pASG-IBA3 expression vector (IBA GmbH, Göttingen, Germany) such that the produced EctC and EctD proteins were attached at their C-termini to a *Strep*-tag II affinity peptide. We refer in the following to these recombinant proteins as (*Nm*)EctC and (*Nm*)EctD.

The (*Nm*)EctC and (*Nm*)EctD proteins were effectively overproduced in *E. coli* BL21 and could be purified to a high degree of homogeneity by affinity chromatography on a Streptactin matrix (Fig. 3). Size exclusion chromatography demonstrated that preparations of both enzymes were essentially free of protein aggregates (Fig. 3). The molecular masses of the (*Nm*)EctC-*Strep*-tag II and (*Nm*)EctD-*Strep*-tag II proteins calculated from their corresponding gene sequences are 15.9 kDa and 35.8 kDa respectively. The two recombinant proteins eluted from the size exclusion columns as species with a calculated molecular mass of 31.9 kDa and 71.4 kDa, respectively, indicating that both the (*Nm*)EctC and (*Nm*)EctD enzymes are dimers in solution (Fig. 3). Such a quaternary assembly has previously also been reported for ectoine synthases (Ono *et al.*, 1999a; Kobus *et al.*, 2015) and ectoine hydroxylases (Höppner *et al.*, 2014; Widderich *et al.*, 2014a) from members of the *Bacteria*.

We then determined a set of basic biochemical parameters for the (*Nm*)EctC and (*Nm*)EctD enzymes and subsequently used these data to set up optimal enzyme activity assays. The (*Nm*)EctC and (*Nm*)EctD enzymes had similar temperature profiles with optima at 30°C and 35°C respectively (Fig. 4A). Both also possessed a similar pH dependency with activity optima of 7.0 and 7.5, for the (*Nm*)EctC and (*Nm*)EctD proteins respectively (Fig. 4B). The activity of the ectoine hydroxylase from '*Ca.* N. maritimus' SCM1 was only slightly activated by increases in ionic strength (elicited either through the addition of NaCl or KCl to the enzyme assay buffer), whereas the activity of the ectoine synthase was restricted to a narrow concentration range of salts (Fig. 4C and D).

Using optimized enzyme assay conditions, we determined the kinetic parameters of the ectoine synthase and of the ectoine hydroxylase using their natural substrates N-γ-L-2,4-acetyl-diaminobutyrate (Ono *et al.*, 1999b) and ectoine (Bursy *et al.*, 2007; Widderich *et al.*, 2014a) respectively. Both enzymes showed Michaelis–Menten-type kinetics. The apparent kinetic parameters for the ectoine synthase were: $K_m = 6.4 \pm 0.6$ mM; $v_{max} = 12.8 \pm 0.4$ U mg^{-1}, $k_{cat} = 5.7$ s^{-1}, $k_{cat}/K_m = 0.9$ s^{-1} mM^{-1} (Fig. 5A). Those for the ectoine hydroxylase were: K_m (ectoine) $= 3.8 \pm 0.5$ mM; K_m (2-oxoglutarate) $= 3.1 \pm 0.6$ mM; $v_{max} = 1.8 \pm 0.1$ U mg^{-1}, $k_{cat} = 1.0$ s^{-1}, $k_{cat}/K_m = 0.3$ s^{-1} mM^{-1} (Fig. 5B and C).

The basic biochemical features and kinetic parameters of the '*Ca.* N. maritimus' SCM1 ectoine hydroxylase are very similar to those of a considerable number of previously characterized EctD proteins (eight enzymes in total) from members of the *Bacteria* (Bursy *et al.*, 2007; 2008; Widderich *et al.*, 2014a). However, this situation is somewhat different for the ectoine synthase, an enzyme for which only two representatives have been kinetically characterized in some detail so far, namely the enzymes from *H. elongata* and *Acidiphilium cryptum* (Ono *et al.*, 1999a; Moritz *et al.*, 2015). The ectoine synthase from *H. elongata* is a highly salt tolerant enzyme (Ono *et al.*, 1999a), whereas both the EctC proteins from *A. cryptum* JF-1 (Moritz *et al.*, 2015) and *Ca.* N. maritimus' SCM1 (Fig. 4C and D) are sensitive to high salt concentrations. Moritz and colleagues (2015) suggested that these different sensitivities of the EctC enzymes from *H. elongata* and *A. cryptum* JF-1 against high salt concentrations result from their different overall net-negative charge (calculated isoelectric points: 4.87 and 6.03 for the EctC proteins from *H. elongata* and *A. cryptum* respectively). The calculated pI of the '*Ca.* N. maritimus' SCM1 EctC protein is 5.65, and hence closer to that of *A. cryptum* JF-1; its salt-sensitive properties (Fig. 4C and D) is thus consistent with the hypothesis put forward by Moritz and colleagues (2015).

The amino acid sequences of key enzymes for ectoine/hydroxyectoine biosynthesis in Archaea cluster with their counterparts from Bacteria

The amino acid sequences of the 723 EctC-type proteins retrieved from the above-described database analysis were aligned and they exhibited, relative to the (*Nm*)EctC protein, a degree of sequence identity ranging between 79% (EctC from the marine Groupe_I Thaumarchaeota SCGC_AAA799-E16) and 33% (EctC from *Micrococcus luteus*). We then used this amino acid sequence alignment to construct a rooted tree of EctC-type proteins using tools provided via the iTOL web server (Letunic and Bork, 2011), in order to assess the taxonomic distribution

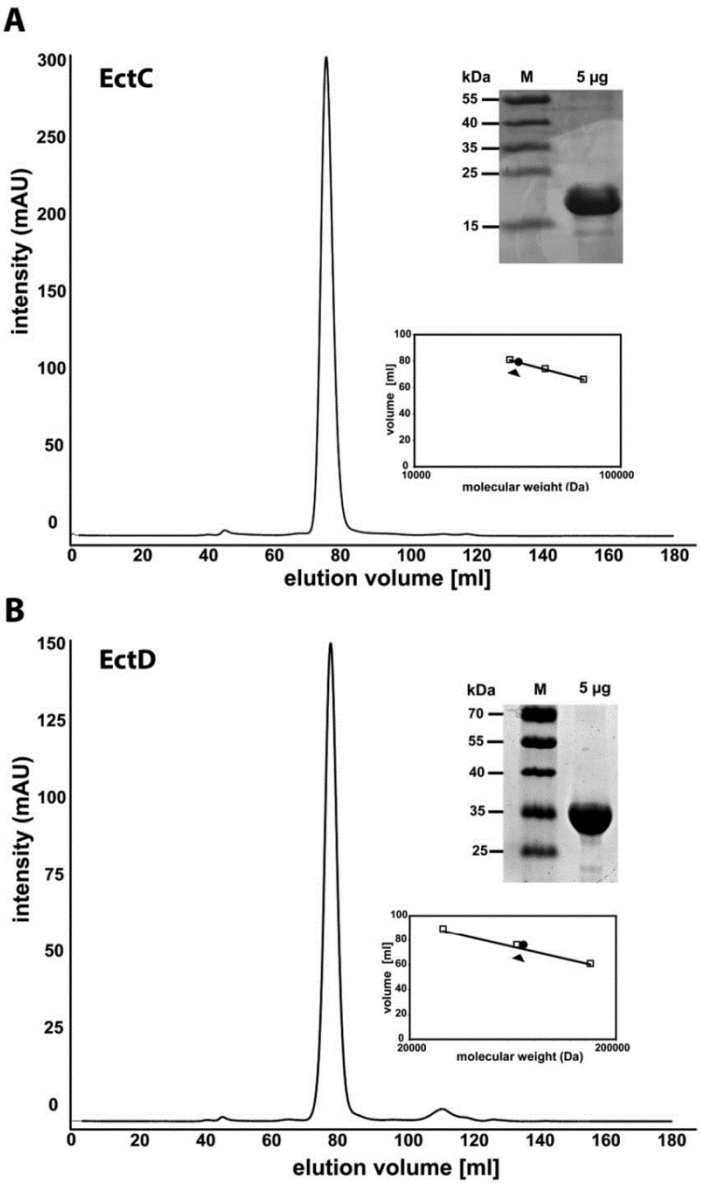


Fig. 3. Assessment of the purity of the (*Nm*)EctC-*Strep*-tag II and (*Nm*)EctD-*Strep*-tag II proteins and their quaternary assembly. (A) the (*Nm*)EctC-*Strep*-tag II and (B) (*Nm*)EctD-*Strep*-tag II recombinant proteins were isolated by affinity chromatography on a Streptactin matrix and analysed by SDS-polyacrylamide gel electrophoresis and size exclusion chromatography. (A) (*Nm*)EctC-*Strep*-tag II and (B) (*Nm*)EctD-*Strep*-tag II. The black arrowhead in the inserts of (A) and (B) indicate the calculated molecular mass of the recombinant proteins relative to that of marker proteins. For the (*Nm*)EctC-*Strep*-tag II protein the following marker proteins were used: albumin (66 kDa), ovalbumin (43 kDa), carboanhydrase (29 kDa). For the (*Nm*)EctD-*Strep*-tag II protein, the following marker proteins were used: alcohol-dehydrogenase (150 kDa), albumin (66 kDa), carboanhydrase (29 kDa).

of the 12 archaeal proteins and their affiliation with the nearest bacterial orthologues (Fig. 6). Nineteen microbial phyla were represented among those 723 microorganisms that possess EctC-type proteins; 16 of these phyla were derived from *Bacteria*, and three were derived from *Archaea*. The few EctC orthologues from archaeal strains cluster at three locations in the EctC-based phylogenetic tree (Fig. 6) (a phylogenetic tree listing the names of the EctC-possessing microorganisms is provided in Fig. S1), suggesting that the corresponding *ect* genes were most likely obtained by lateral gene transfer events from bacterial donor strains harbouring closely related orthologues.

In the eight members of the Thaumarchaeota possessing *ect* gene clusters, all EctC orthologues form a monophyletic clade by themselves. EctC proteins of mostly *Halomonas* species and related genera of the *Halomonadaceae*, which are affiliated with the

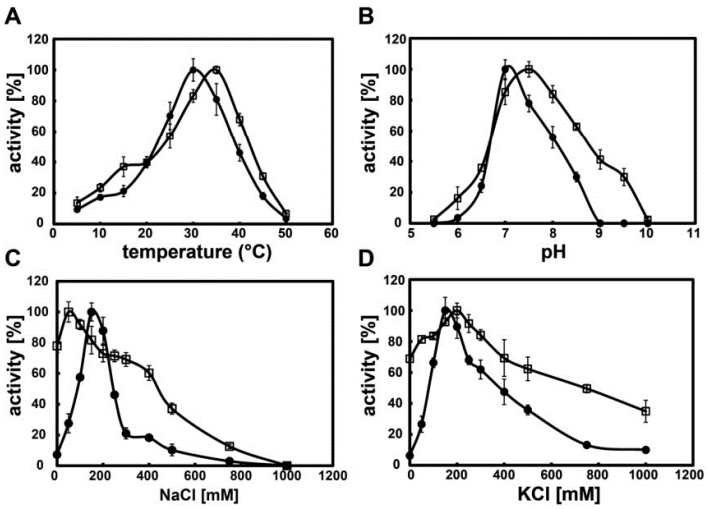


Fig. 4. Biochemical properties of the (*Nm*)EctC-*Strep*-tag II and (*Nm*)EctD-*Strep*-tag II proteins. Black dots: (*Nm*)EctC-*Strep*-tag II; open squares: (*Nm*)EctD-*Strep*-tag II.

Gammaproteobacteria, represent the most similar adjacent clade. Because all microorganisms from these two clades populate marine habitats, lateral transfer of the *ect* genes from a member of the *Halomonadaceae* to an ancestral strain of the genus *Nitrosopumilus* is plausible. The presence of some EctC sequences from marine *Alphaproteobacteria* (e.g. genera *Nesiotobacter*, *Zhangella* or *Marinitalea*) within the *Halomonadaceae* clade suggests that lateral gene transfer events have still been ongoing after the *ect* genes were transferred to an ancestor of *Nitrosopumilus*. The acquisition of ectoine/hydroxyectoine biosynthetic genes in the indicated strains by a lateral gene transfer event(s) is further supported by the fact that none of the AOA [*Nitrosopumilus sediminis* AR2, *Nitrosopumilus salaria* BD31, *Nitrosopumilus koreensis* MY1, *Nitrosoarchaeum limnia* (strains SFB1 and BG20), *Nitrosotenuis uzonensis* N4, *Nitrosophera gargensis* and *Nitrosocaldus* sp.] contain ectoine/hydroxyectoine biosynthetic genes.

The other two clades of archaeal EctC orthologues are present in some members of the methanogenic genera *Methanothrix* and *Methanobacterium*, respectively, but they are not found in all sequenced strains of these genera. In case of *Methanothrix*, the closest relatives are from the strictly anaerobic sulfate-reducing deltaproteobacterial genus *Desulfarculus* and the aerobic alphaproteobacterial genus *Sneathiella*, and they form a common clade. Although very diverse in their taxonomy, all these strains were isolated from freshwater sediments or wastewater treatment plants, creating ample opportunity for lateral gene transfer. From the large phylogenetic distance between *Nitrosopumilus* and *Methanothrix*, it is clear that the respective gene transfer events must have happened independently. Among the seven genome-sequenced species of the genus *Methanobacterium*, *ect* genes appear only to be present in the species *M. formicicum*. The respective EctC sequences form a common cluster with EctC orthologues from terrestrial and marine *Bacillaceae*, but also some *Betaproteobacteria* and *Gammaproteobacteria*. Therefore, it may be inferred that the *ect* genes in *M. formicicum* have been acquired by a very recent lateral transfer from one of the bacterial species with closely related *ectABC* sequences.

Since it is highly likely that the *ect* gene clusters present in members of the *Archaea* (Fig. 1B) were acquired via lateral gene transfer events, we attempted to identify possible bacterial donors by using the corresponding amino acid sequences of the ectoine/hydroxyectoine biosynthetic enzymes as query sequences in a BLASTP search of microbial genomes (Nordberg *et al.*, 2013) and then recording the top 10 hits. As documented in Table S1 as the results for this type of analysis for the Thaumarchaeota '*Ca.* N. maritimus' SCM1, '*Ca.* *Nitrosopumilus*' sp. SJ, '*Ca.* Nitrosopumilus koreensis' AR1 and '*Ca.* Nitrosolumilus' sp. AR2, no clear pattern emerged that would allow to precisely pinpoint possible bacterial donor species for the archaeal *ect* gene clusters.

Taken together, our phylogenetic assessment of EctC-type proteins is consistent with our biochemical and kinetic studies of the '*Ca.* N. maritimus' SCM1 ectoine synthase (Figs 4 and 5) that ascribe bacterial-like properties to this archaeal protein. In analysing the rooted phylogenetic tree of EctC-type proteins, we noted that all orphan EctC-like proteins (Kurz *et al.*, 2010; Widderich *et al.*, 2014a) cluster together and form a distinct branch close to the root of the tree (Fig. 6). Among the 723 EctC-type proteins analysed, 24 belong to this group, and none was from an archaeon. The taxonomic affiliation

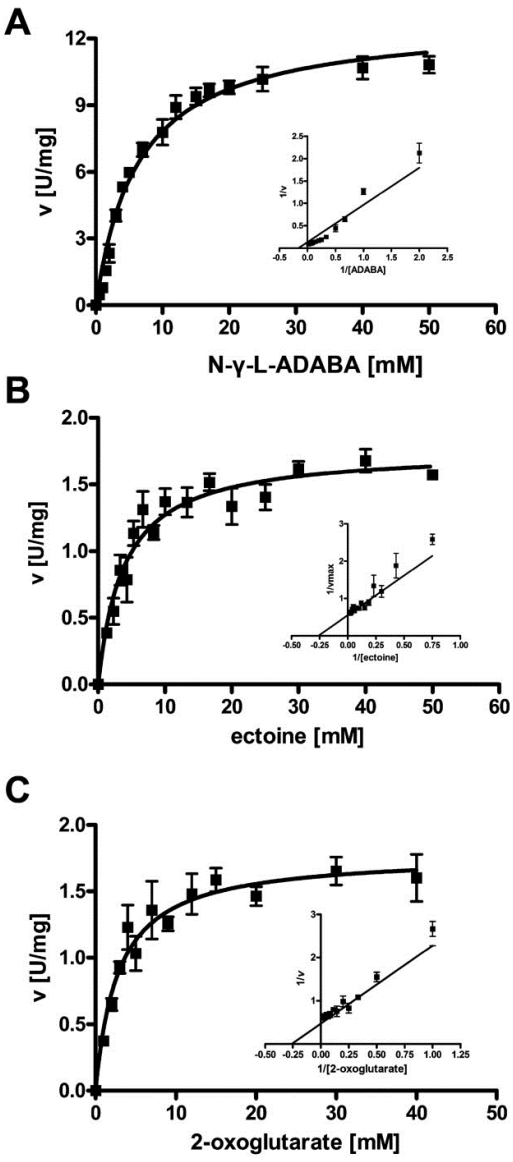


Fig. 5. Kinetic parameters of the (A) EctC for its substrate Nγ-ADABA and of the ectoine hydroxylase EctD for its substrate ectoine (B) and its co-substrate 2-oxoglutarate (C).

harbouring these orphan *ectC* genes is rather diverse (Fig. 6).

A novel transcriptional organization of the ect gene cluster in 'Ca. N. maritimus' SCM1

Inspection of the organization of the *ectABCD* gene cluster of '*Ca.* N. maritimus' SCM1 and those of other members of this genus (Fig. 1B) revealed a conserved genetic arrangement that is also found in many members of the *Bacteria* (Widderich *et al.*, 2014a). The distance between the *ectA* and *ectB* genes is three bp, three bp between *ectB* and *ectC*, and the *ectC* and *ectD* coding regions overlap by two bp. These observations suggest that the *ectABCD* gene cluster of '*Ca.* N. maritimus' SCM1 is transcribed as part of an operon. While the genetic organization of the 3′ region of the *ect* genes is variable (Fig. 1B), they are all preceded in the *Nitrosopumilus* species and other marine Thaumarchaeota by an open reading frame (hyp) which codes for a hypothetical protein (Hyp; 125 amino acids) of unknown function. The coding region of the *hyp* gene and *ectA* in '*Ca.* N. maritimus' SCM1 overlaps by three bp, suggesting that the *hyp* gene is co-transcribed with the *ectABCD* gene cluster (Fig. 1B). As a specialty among members the genus *Nitrosopumilus* and other marine *Thaumarchaea* containing *ect* genes, the *ectABCD* genes of '*Ca.* N. maritimus' SCM1 are followed by an open reading frame (Fig. 1B) that was originally annotated to code for a protein of unknown function (Walker *et al.*, 2010). Our renewed database searches identified this protein as a member of the MscS family (see below), mechanosensitive channels that serve as safety valves against the cell-disrupting consequences of severe osmotic downshifts (Levina *et al.*, 1999). They are ubiquitously found in both *Bacteria* and *Archaea* (Kloda and Martinac, 2002; Booth and Blount, 2012; Naismith and Booth, 2012; Wilson *et al.*, 2013; Booth, 2014; Booth *et al.*, 2015). We refer to this protein in the following as the (*Nm*)MscS channel. The distance between the end of the *ectD* gene and the start of the *mscS* gene is 25 bp (Fig. 1B), suggesting that the *mscS* gene might be co-transcribed with the *ectABCD* gene cluster as well. Overall, the tight physical organization of the *hyp-ectABCD-mscS* gene cluster (Gene IDs: Nmar_1347, Nmar_1346, Nmar_1345, Nmar_1344, Nmar_1343, Nmar_1342) suggests that these genes are transcribed as an operon.

To test for a possible co-transcription of the *hyp-ectABCD-mscS* genes (Fig. 7A), we carried out an reverse transcription polymerase chain reaction (RT-PCR) analysis using RNA samples that were isolated from cells of '*Ca.* N. maritimus' SCM1 that had been grown under increased osmotic stress conditions (48 g NaCl l^{-1} which corresponds to 0.94 M NaCl) (Fig. 2A). The data presented in Fig. 7B demonstrate that the *hyp-ectABCD-mscS* gene cluster is indeed expressed as a transcriptional unit. Co-transcription of ectoine/hydroxyectoine biosynthetic genes (Widderich *et al.*, 2014a) has been routinely observed in members of the *Bacteria* (Kuhlmann and Bremer, 2002; Bursy *et al.*, 2007; Kuhlmann *et al.*, 2008). However, there is no report where they are co-transcribed with a gene for a mechanosensitive channel.

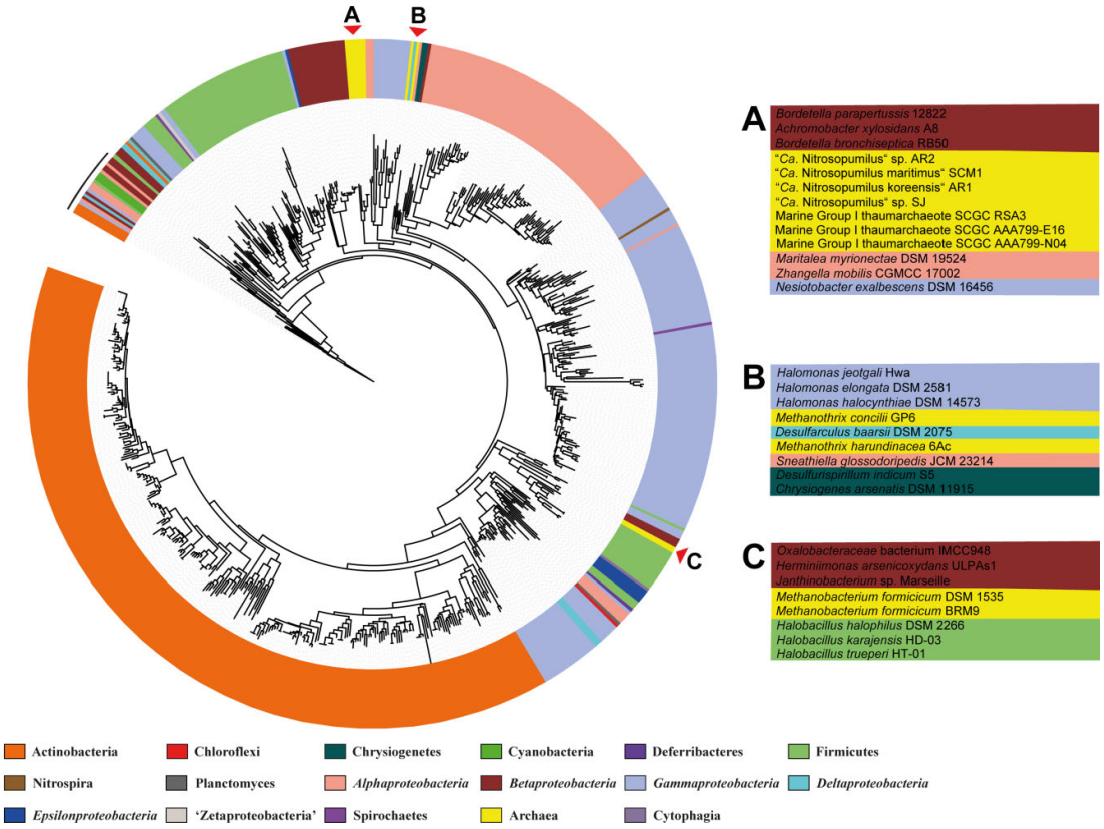


Fig. 6. Phylogenetic analysis of EctC-type proteins. Based on an amino acid sequence alignment of 723 EctC-type proteins, a rooted phylogenetic tree was constructed with the ITOL program (Letunic and Bork, 2011). The different phyla are marked by different colours and, for simplicity, the names of the *ectC*-possessing microorganisms were left out. An expanded version of the tree containing the names of the microorganisms predicted to produce the EctC protein is given in Fig. S1. The three regions in the phylogenetic tree populated by archaeal EctC proteins are highlighted by red arrowheads. An enlarged section of the phylogenetic tree representing the archaeal EctC proteins and their nearest bacterial orthologues are shown in A to C. The position of those 24 EctC-type proteins that originate from bacteria that lack identifiable *ectAB* genes are marked above the phylogenetic tree by a black bar.

The eight *ect* gene clusters present in Thaumarchaeota are all preceded by a gene (*hyp*) (Fig. 1B) that encodes a protein (125 amino acids) of unknown function. These Hyp proteins possess a degree of amino acid sequence identity ranging between 82% to 31%, and such proteins were found through database searches also in the Thaumarchaeota MY2 and N4, both of which do not possess *ect* genes. The Hyp proteins show no significant sequence similarity to annotated proteins in databases, and we can thus neither speculate about their potential function, nor can we comment on the obvious question why the *hyp* gene is co-transcribed with the *ect* gene cluster in various Thaumarchaeota (Fig. 7).

Expression of the *ect* genes in *Bacteria* is typically induced when the cells are exposed either to sustained or suddenly imposed osmotic stress (Kuhlmann and Bremer, 2002; Calderon *et al.*, 2004; Bursy *et al.*, 2007; Kuhlmann *et al.*, 2008; Salvador *et al.*, 2015). We therefore wondered whether this was also the case in the archaeon '*Ca.* N. maritimus' SCM 1. We therefore isolated total RNA from cells that had been grown under low, optimal and high salt conditions (Fig. 2A) (13, 25 and 39 g NaCl l^{-1}, respectively) and assed the transcript levels of a gene positioned either early (*ectA*) or late (*mscS*) in the *hyp-ectABCD-mscS* poly-cistronic messenger (m)RNA via q-RT-PCR. The cells of the cultures grown under different salt-stress conditions were all harvested when they were in the same growth phase to exclude possible effects of this parameter on the quantitative RT-PCR data. We used the transcript of the *amoA* gene (Gene ID: Nmar_1500) of '*Ca.* N. maritimus' SCM 1, as a reference (Nakagawa and Stahl, 2013) to benchmark the *ectA* and *mscS* transcript

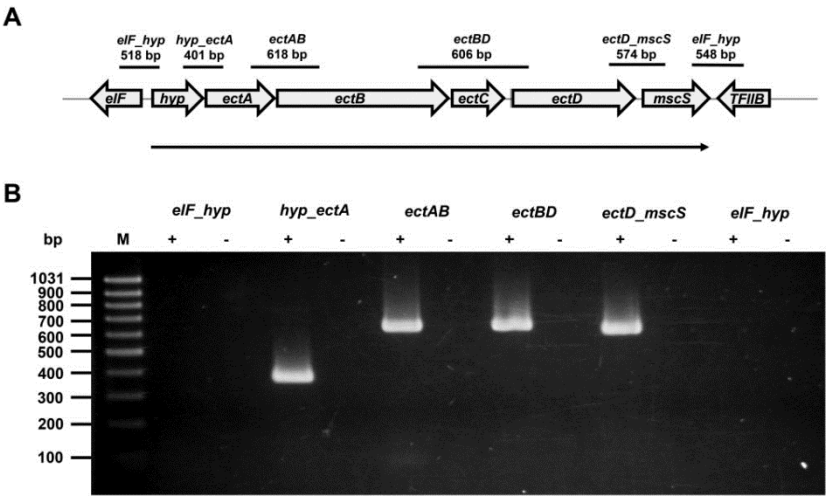


Fig. 7. Analysis of the co-transcription of the *hyp-ectABCD-mscS* gene cluster from '*Ca.* N. maritimus' SCM1 by RT-PCR.
A. Genetic organization of the *hyp-ectABCD-mscS* gene cluster and predictions of the size of the PCR fragments made under the assumption that a given set of genes is co-transcribed.
B. Analysis of the sizes of the DNA fragments generated in the RT-PCR experiment by agarose gel electrophoresis. The symbol '+' indicates those samples reacted with reverse transcriptase, and the symbol '-' denotes samples prepared in its absence to ensure that the observed PCR products did not result from DNA contaminations of the RNA samples.

levels. Indeed, the *amoA* transcript turned out to be a suitable reference marker for salt-stressed and non-stressed cells since the level of its mRNA did not vary significantly under the above-described growth conditions (Fig. 2C). In contrast, the level of the *ectA* and of the *mscS* transcripts was strongly upregulated from an already substantial level in response to salt stress; by 5.7-fold and 5.1-fold respectively (Fig. 2C). An even greater enhanced induction of the *ectA* and *mscS* transcript levels was observed when '*Ca.* N. maritimus' SCM1 cells that had been grown under optimal salt conditions (Fig. 2A) (26 g NaCl l^{-1}) were subjected to a severe osmotic upshift (by adding 23 g NaCl l^{-1} to the culture). This severe salt shock triggered a 52-fold and 76-fold increase in the *ectA* and *mscS* transcripts from their levels in cells grown under optimal salt conditions (Fig. 2D). Hence, both chronic and acute salt stresses are important environmental cues leading to strongly enhanced transcription of the *hyp-ectABCD-mscS* operon.

(Nm)MscS is a functional mechanosensitive channel

Since there is no report in the literature describing the co-transcription of an *ect* gene cluster with a gene encoding a mechanosensitive channel, we wondered whether the (*Nm*)MscS channel was functional. To study the functionality of the (*Nm*)MscS channel, we used a genetic complementation experiment (Levina *et al.*, 1999) with an *E. coli* mutant strain (MJF641) (also referred to as the Δ7 strain) that lacks all currently know MscM-, MscS- and MscL-type mechanosensitive channels (Edwards *et al.*, 2012). We obtained a synthetic version of the '*Ca.* N. maritimus' SCM 1 *mscS* gene that had been optimized for its expression in *E. coli* and placed it under the transcriptional control of the *lac* promoter present on plasmid pTrc99a (Amann *et al.*, 1988), thereby yielding plasmid pLC18. We then used a test developed by Levina and colleagues (1999) to assess the functionality of mechanosensitive channels *in vivo*. It relies on a rapid down-shock of high osmolarity grown cells lacking or expressing channel-forming proteins and then monitoring cell viability (Levina *et al.*, 1999). The parent (strain Frag1) of the Δ7 mutant-strain carrying the vector plasmid pTrc99a survived such an osmotic downshift essentially unscathed, whereas only 12.9% (±5%) of the cells of the Δ7 mutant-strain MJF641 (pTrc99a) survived such a treatment. In contrast, when MJF641 (pLC18; *mscS*$^{+}$) was subjected to such an osmotic down-shock, the (*Nm*)MscS protein rescued cellular survival to a large extent (62% ± 5% surviving cells) (Fig. 8A). This experiment therefore unambiguously demonstrates that the gene following the *ectABCD* gene cluster (Figs 1B and 7A) in the '*Ca.* N. maritimus' SCM1 genome (Walker *et al.*, 2010) encodes a functional mechanosensitive channel.

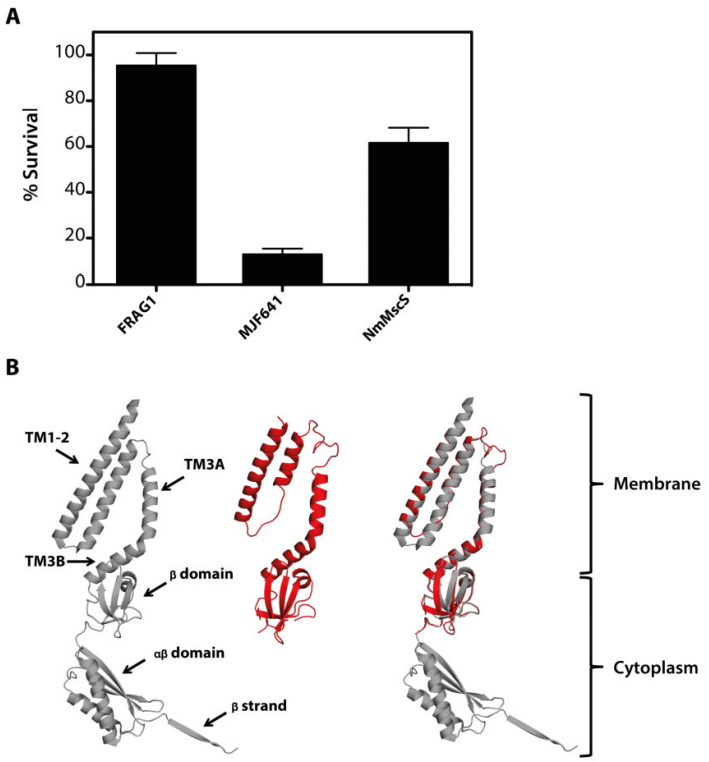


Fig. 8. Functional assessment of the (*Nm*)MscS protein and *in silico* prediction of its structure.
A. Cells of the *E. coli* strain MJF641 which is defective in all currently known MscL-, MscS- and MscM-type channels (also referred to as Δ7 mutant) (Edwards *et al.*, 2012) and its parent strain Frag1 carrying the empty vector plasmid pTrc99a (Amann *et al.*, 1988) were grown in high-salinity minimal medium (with 0.3 M NaCl) and were then subjected to a rapid osmotic downshift (minimal medium without additional NaCl) (Levina *et al.*, 1999). The same procedure was used with strain MJF641 harbouring a plasmid (pLC18) that contains the '*Ca.* N. maritimus' SCM1 *mscS*$^{+}$ gene expressed under the control of the *lac* promoter present on pTrc99a. The number of cells surviving such rapid osmotic downshifts was determined by spotting them on LB agar plates. The data shown represent the mean and standard deviation of four independently conducted experiments.
B. Crystal structure of a monomer of the *E. coli* MscS protein (left) from the open *E. coli* MscS heptameric channel assembly (protein database accession code: 2VV5) (Wang *et al.*, 2008) was automatically chosen by the SWISS-MODEL web server (Biasini *et al.*, 2014) as the template to generate an *in silico* model (middle) of the (*Nm*)MscS protein. The PYMOL suite (https://www.pymol.org/) was used to visualize these structures and overlay them with each other (right).

In silico analysis of the (Nm)MscS mechanosensitive channel

Mechanosensitive channels acting as safety valves against osmotic downshifts can be grouped into three types (MscM, MscS, MscL) and provide the microbial cell with a graded stress response to an osmotic challenge through their different threshold levels of opening and their different diameters of the fully open channels (Haswell *et al.*, 2011; Booth and Blount, 2012; Edwards *et al.*, 2012; Booth, 2014). A database search identified the '*Ca.* N. maritimus' SCM1 protein as a member of the MscS family with an overall level of amino acid sequence identity to the *E. coli* MscS protein (Levina *et al.*, 1999) of about 12%. However, in comparison to the *E. coli* protein, (*Nm*)MscS lacks about 100 amino acids at its carboxy-terminus (Fig. S2). When this is taken into consideration, and a stretch of badly aligning 20 amino acids from the N-terminus are removed as well from the comparison of the two proteins, the degree of amino acid sequence identity increases to about 16%. Consistent with this amino acid sequence alignment, the SWISS MODEL server (http://swissmodel. Expacy.org/) (Biasini *et al.*, 2014) automatically chose a monomer of the homo-heptamertic *E. coli* MscS channel in its 'open' form (PDB entry 2VV5) (Wang *et al.*, 2008) as its modelling template (Fig. 8B). Like its *E. coli* MscS counterpart (Bass *et al.*, 2002; Wang *et al.*, 2008), the (*Nm*)MscS protein is predicted to contain three membrane-spanning segments and a carboxy-terminus that protrudes into the cytoplasm (Fig. 8B). Inspection of the (*Nm*)MscS *in silico* model revealed an important difference with respect to the *E. coli* MscS protein, as the cytoplasmic αβ-domain and an adjacent β-sheet is missing (Fig. 8B).

Interestingly, a protein similar to that of (*Nm*)MscS (Fig. S2) is encoded in the genome of '*Ca.* Nitrosopumilus' sp. AR2. However, in contrast to the situation in '*Ca.* N. maritimus' SCM1, the putative *mscS* gene is transcribed divergently from the *hyp-ectABCD* gene cluster (Fig. 1B). Of note is also that the genome sequence of '*Ca.* N. maritimus' SCM1 (Walker *et al.*, 2010) encodes in addition to the (*Nm*)MscS protein (gene ID: Nmar_1342), four other MscS-encoding genes (gene ID: Nmar_1337. Nmar_1773, Nmar_1335, Nmar_0971). These MscS-type proteins vary considerably with respect to the length of their carboxy-termini. '*Ca.* Nitrosopumilus maritimus' SCM1 does not possess a gene for an MscL-type channel.

We also inspected the genome sequences of all other archaeal strains that possess *ect* gene clusters for the presence of *mscS* and *mscL*-type genes. Each of these microorganisms possesses multiple numbers (between one and four) of *mscS*-type genes but only *M. harundinacea* 6Ac and *Methanotrix concilii* GP-6 possess *mscL*-type genes as well (Table S2). Hence, MscS-type mechanosensitive channels (Levina *et al.*, 1999; Wang *et al.*, 2008) seem to be the preferred device to physiologically cope with rapid osmotic downshifts by those *Archaea* whose genome sequences we have inspected.

Discussion

Ectoine and its derivative 5-hydroxyectoine are well-recognized compatible solutes that are synthesized, or taken up, in response to osmotic stress by many microorganisms (Pastor *et al.*, 2010; Widderich *et al.*, 2014a). Their biophysical properties allow them not only to serve as effective stress protectants but also to function as chemical chaperones as they are able to preserve the functionality of proteins, macromolecular complexes, membranes and even entire cells (Lippert and Galinski, 1992; Manzanera *et al.*, 2002; Harishchandra *et al.*, 2010; Pastor *et al.*, 2010; Kunte *et al.*, 2014; Tanne *et al.*, 2014). Our updated database search identified the signature enzyme for ectoine production, the ectoine synthase (Ono *et al.*, 1999b; Widderich *et al.*, 2014a), in about 4.7% of 27.789 microbial genome sequences deposited in the JGI database at the time of the search. Predicted ectoine producers are found in 19 microbial phyla. Only three of these represent *Archaea*, and only 12 archaeal strains belonging to the genera *Nitrosopumilus*, *Methanothrix* and *Methanobacterium* are predicted to engage in ectoine synthesis among the 557 *Archaea* with deposited genome sequences. This data set therefore reinforces our previous conclusion that ectoine is a compatible solute primarily synthesized by members of the *Bacteria* (Widderich *et al.*, 2014a).

It is notable that among the 12 archaeal strains possessing ectoine biosynthetic genes, eight are affiliated with the genus *Nitrosopumilus*. We therefore used '*Ca.* N. maritimus' SCM1 (Könneke *et al.*, 2005; Walker *et al.*, 2010) as a model system to study ectoine/hydroxyectoine synthesis and the genetic control of its *ect* gene cluster in this physiologically well-characterized and globally abundant member of the domain *Archaea* (Stahl and de la Torre, 2012; Offre *et al.*, 2013). Both the transcription of the ectoine/hydroxyectoine biosynthetic genes and the production of these compatible solutes in '*Ca.* N. maritimus' SCM1 are responsive to osmotic stress (Fig. 2), thereby strongly suggesting that ectoines serve an osmostress-protective function in this Thaumarchaeon, as firmly established in many members of the *Bacteria* (Pastor *et al.*, 2010; Kunte *et al.*, 2014). We surmise that the data presented here on these issues for '*Ca.* N. maritimus' SCM1 are relevant for an understanding of the physiology and genetics of the production of these compatible solutes by other strains and species of the Thaumarchaeota, and we suggest that is also true for those species of the methanogens *Methanothrix* and *Methanobacterium* carrying ectoine biosynthetic genes (Fig. 1B).

A phylogenetic tree derived from an amino acid sequence alignment of 723 unique EctC-type proteins revealed that the 12 archaeal EctC orthologues cluster in three separate locations among their bacterial counterparts (Figs 6 and S1). These data strongly suggest that the *Archaea* harbouring *ectABC(D)* gene clusters have acquired them via horizontal gene transfer events from bacterial donor strains living in the same habitats. This suggestion is also consistent with our finding that the key enzymes for ectoine biosynthesis, the ectoine synthase (EctC) and the ectoine hydroxylase (EctD) from '*Ca.* N. maritimus' SCM1 possessed biochemical and kinetic properties (Figs 4 and 5) resembling those of their bacterial counterparts (Ono *et al.*, 1999b; Bursy *et al.*, 2007; 2008; Widderich *et al.*, 2014a; Moritz *et al.*, 2015). Since the 12 EctC proteins derived from *Archaea* are found in three different locations in the ectoine synthase-based phylogenetic tree (Figs 6 and S1), it seems likely that the *ectABC(D)* genes were introduced into members of the genera *Nitrosopumilus*, *Methanothrix* and *Methanobacterium* by separate events. However, by comparing the amino acid sequences of the archaeal ectoine/hydroxyectoine biosynthetic enzymes to their nearest orthologues from *Bacteria* (Table S1), we were unable to derive at a consistent picture of the potential donor *Bacteria*.

It is well known that horizontal gene transfer is an important driver of microbial evolution helping microorganisms to rapidly obtain new metabolic capabilities and develop new stress management skills so that they can better cope with the situation in their original habitat, or explore new ecosystems (Ochman *et al.*, 2000; Soucy *et al.*, 2015). This process shaped the gene content of archaeal genomes in a major way (Nelson-Sathi *et al.*, 2015). For instance, it was concluded from a metagenomic analysis of deep Mediterranean waters that about 24% of the genes derived from planktonic Thaumarchaeota were of bacterial origin (Deschamps *et al.*, 2014).

While this paper was under review, the genome sequences of two novel *Nitrosopumilus* strains originating from coastal surface waters of the Northern Adriatic Sea became available. We found that neither '*Ca.* Nitrosopumilus piranensis' D3C, nor '*Ca.* Nitrosopumilus adriaticus' NF 5 (Bayer *et al.*, 2015) possess ectoine

biosynthetic genes. Remarkably, based upon a 16 S-23S rRNA tree derived from cultivated Thaumarchaeota with sequenced genomes, '*Ca.* Nitrosopumilus piranensis' D3C, is the closest relative of '*Ca.* N. maritimus' SCM1 and '*Ca.* Nitrosopumilus koreensis' AR1 (Bayer *et al.*, 2015), both of which possess ectoine/hydroxyectoine biosynthetic gene clusters (Fig. 1B). Hence, it seems likely that the possession, or lack, of ectoine/hydroxyectoine biosynthetic genes among the AOA reflects specific ecological niche adaptive evolutionary processes (Blainey *et al.*, 2011; Gubry-Rangin *et al.*, 2015).

No genetic system is currently available for '*Ca.* N. maritimus' SCM1 (Walker *et al.*, 2010) that would allow the construction of a mutant derivative lacking the *ectABCD* gene cluster. However, the isolation of a taxonomically very close relative, '*Ca.* Nitrosopumilus piranensis' D3C, that lacks these genes naturally (Bayer *et al.*, 2015) provides the opportunity to compare the physiology and growth properties of these two Thaumarchaeota to reveal possible benefits of ectoine/hydroxyectoine production.

The function-preserving features of ectoines have triggered considerable biotechnological interest and led to the development of an industrial scale production process that relies on a severe osmotic down-shock to release newly synthesized ectoines from microbial producer cells (Schwibbert *et al.*, 2011; Kunte *et al.*, 2014). This so-called bacterial milking procedure builds on an evolutionarily conserved cellular emergency stress reaction to sudden osmotic down-shocks that relies on the transient opening of MscM-, MscS- and MscL-type mechanosensitive channels to prevent cell rupture (Booth, 2014). These types of safety valves have a deep evolutionary origin and are found both in *Archaea* and *Bacteria* (Kloda and Martinac, 2002; Pivetti *et al.*, 2003).

Most interestingly, we discovered that the *ectABCD* gene cluster of '*Ca.* N. maritimus' SCM1 is co-transcribed with the gene for an MscS-type mechanosensitive channel (*hyp-ectABCD-mscS*) (Fig. 7). While nothing can currently be concluded about the role of the *hyp* gene product, the *mscS*-encoded mechanosensitive channel is functional (Fig. 8A). A genetic configuration that entails the co-transcription of a *mscS* gene along with the *ectABC(D)* operon has not been described before in any ectoine/hydroxyectoine-producing microorganism. The *hyp-ectABCD-mscS* operon seems to be a rather sophisticated genetic device to counteract the detrimental effects of osmotic fluctuations on cell physiology. The osmotic induction of its transcription allows '*Ca.* N. maritimus' SCM1 to cope with increased salinity through enhanced production of the osmostress protectants ectoine and hydroxyectoine. At the same time, it prepares the high osmolarity stressed cells for an osmotic down-shift that eventually will occur in narrow transition zones from marine to estuarine ecosystems through enhanced preemptive synthesis of the (*Nm*)MscS (Santoro *et al.*, 2008) mechanosensitive channel. We are aware that in the hyperthermophilic Creanarchaeon *Thermoproteus tenax*, an *mscS*-type gene is co-transcribed with enzymes required for the synthesis of the compatible solute trehalose (Zaparty *et al.*, 2013). However, in contrast to the biological data we provide here for the (*Nm*)MscS protein (Fig. 8A), the functionality of the Msc protein from *T. tenax* has not yet been experimentally assessed.

Both the MscS and the MscL mechanosensitive channels are widely considered as non-specific pores when they are fully opened and they have calculated pore diameters of about 13 Å and 30 Å respectively (Cruickshank *et al.*, 1997; Wang *et al.*, 2008). However, the fact that both in '*Ca.* N. maritimus' SCM1 and in *T. tenax*, *mscS*-type genes are co-expressed with genes encoding enzymes for the synthesis of the compatible solutes ectoine/hydroxyectoine and trehalose, respectively, raises the question if the (*Nm*)MscS and (*Tt*)MscS channels are not endowed with a certain degree of substrate specificity. It is of interest to consider in this context the properties of the MscCG mechanosensitive channel of *Corynebacterium glutamicum*. Like (*Nm*)MscS, MscCG is a MscS-type channel and participates in the osmotic stress response of *C. glutamicum* (Börngen *et al.*, 2010). However, it also serves as an export channel for glutamate *in vivo* (Nakamura *et al.*, 2007; Becker and Krämer, 2015). Despite these latter observations, electrophysiological studies did not suggest that the MscCG channel possesses substrate specificity for glutamate (Becker *et al.*, 2013; Nakayama *et al.*, 2013).

Most members of the diverse MscS family are structurally built on a common core architecture in which the monomer possesses three trans-membrane-spanning segments. The membrane-embedded helix one and two form the sensor for membrane tension and part of the broken helix 3 lines the pore of the heptameric channel assembly (Bass *et al.*, 2002; Wang *et al.*, 2008; Naismith and Booth, 2012) (Fig. 8B). This latter helix provides a structural link between the pore domain and the carboxy-terminal part of the MscS monomer that forms an extended funnel-like vestibule inside the cytoplasm in the fully assembled heptameric MscS channel. Solutes and solvents must pass on their way out of the cell through this structurally complex vestibule when the channel opens (Bass *et al.*, 2002; Wang *et al.*, 2008; Naismith and Booth, 2012; Wilson *et al.*, 2013; Booth *et al.*, 2015). Through conformational changes, the cytoplasmic cage probably plays an important role for the opening and closing reactions of the MscS channel as a whole. It might also serve as a sensor for increases in macromolecular crowding of the cytoplasm (Rowe *et al.*, 2014), a process that is

strongly influenced by osmotically instigated water fluxes in or out of the cell (Wood, 2011).

We therefore note with considerable interest that (*Nm*)MscS protein lacks, in comparison with the *E. coli* MscS protein (Bass *et al.*, 2002; Wang *et al.*, 2008), a large part of the carboxy-terminus (Fig. S2). This should lead to a structurally rather different make-up of the cytoplasmic portion of the (*Nm*)MscS channel (Fig. 8B). Mechanosensitive channels open in response to increased tension in the plain of the lipid bilayer of the cytoplasmic membrane that results from an increase in turgor (Naismith and Booth, 2012; Booth, 2014). These types of proteins therefore have to closely interact with lipids to perform their physiological function. The types of lipids that are present in the Thaumarchaeon '*Ca.* N. maritimus' SCM1 are very different from those of *E. coli* (Elling *et al.*, 2014), yet the (*Nm*)MscS protein is physiological active in the surrogate *E. coli* host strain (Fig. 8A). Based on the discussed properties of the (*Nm*)MscS protein, it might be well worthwhile to study its electrophysiological properties, gating behaviour and potential substrate specificity in its own right.

Experimental procedures

Chemicals and synthesis of the substrate of the ectoine synthase

Ectoine and hydroxyectoine were kind gifts from Dr Irina Bagyan (bitop AG, Witten, Germany). Anhydrotetracycline, desthiobiotine and the strepavidin affinity matrix for the purification of *Strep*-tag II-labelled protein were purchased from IBA GmbH (Göttingen, Germany). Alkaline hydrolysis of ectoine was performed as reported previously (Kunte *et al.*, 1993) to obtain the substrate, *N*-γ-acetyl-L-2,4-diaminobutyrate (N-γ-ADABA), for the EctC (Peters *et al.*, 1990; Ono *et al.*, 1999b). All chemicals required for this synthetic process were purchased from Sigma Aldrich (Steinheim, Germany) or Acros (Geel, Belgium). Briefly, hydrolysis of ectoine (284 mg, 2.0 mmol) was accomplished in aqueous potassium hydroxide (KOH) (50 ml, 0.1 M) for 20 h at 50°C (Kunte *et al.*, 1993). The reaction mixture was subsequently neutralized with perchloric acid (60% in water, 4 ml), and the precipitated potassium perchlorate was filtered off. The filtrate was concentrated under reduced pressure, and the residue was purified by repeated chromatography on a silica gel column (Merck silica gel 60). Chromatography was performed by using a gradient of ethanol/25% ammonia/water 50:1:2 – 10:1:2 as the eluent to yield N-γ-ADABA. We recovered 192 mg (1.20 mmol; 60%) of N-γ-ADABA from the starting material (284 mg ectoine, 2.0 mmol). The identity and purity of the isolated N-γ-ADABA was unequivocally established both by thin-layer chromatography (TLC) (silica gel 60 F254 TLC plates; Merck) and nuclear magnetic resonance (^{1}H-NMR and ^{13}C-NMR) spectroscopy (Peters *et al.*, 1990; Ono *et al.*, 1998) on a Bruker AVIII-400 or DRX-500 NMR spectrometer. (i) TLC: R_f of N-γ-ADABA = 0.55 (ethanol/25% ammonia/water 7:1:2); (ii) ^{1}H-NMR (400 MHz, D_2O): δ = 3.71 (dd, ^3J(H,H) = 7.6 Hz, ^3J(H,H) = 5.6 Hz, 1H, CH), 3.41 – 3.24 (m, 2H, CH_2), 2.15 – 2.01 (m, 2H, CH_2), 1.99 (s, 3H, CH_3) ppm; (iii) ^{13}C-NMR (100 MHz, D_2O): δ = 177.5 (CO), 177.0 (COOH), 55.3 (CH), 38.3 (CH_2), 33.0 (CH_2), 24.6 (CH_3) ppm.

Recombinant DNA procedures and construction of plasmids

The DNA sequences of the *ectC*, *ectD* and *mscS* genes were retrieved from the genome sequence (accession number: NC_010085) (Walker *et al.*, 2010) of '*Ca.* N. maritimus' SCM1 and were used as templates for the synthesis of codon-optimized versions of these genes for their expression in *E. coli*. These synthetic genes were constructed either by GenScript (Piscataway, USA), or by LifeTechnologies (Darmstadt, Germany). Their DNA sequences were deposited in the NCBI database under accession numbers KR002039 (*ectC*), JN019033 (*ectD*) and KT313590 (*mscS*). To allow the overproduction and purification of the '*Ca.* N. maritimus' SCM1 EctC and EctD proteins in *E. coli*, we constructed recombinant versions of the corresponding genes. The *ectC* and *ectD* sequences were retrieved from the plasmids provided by the suppliers of the synthetic constructs, and inserted into the expression vector pASG-IBA3 (IBA GmbH, Göttingen, Germany). The resulting expression vectors [pWN48 (*ectC*$^+$) and pMP45 (*ectD*$^+$)] permit the effective synthesis of the '*Ca.* N. maritimus' SCM1 recombinant EctC and EctD proteins in *E. coli*. Both proteins are fused at their carboxy-termini with a *Strep*-tag II affinity peptide (NWSHPQFEK). This allows the purification of the (*Nm*)EctC-*Strep*-tag II and (*Nm*)EctD-*Strep*-tag II proteins by affinity chromatography on a streptavidin matrix.

For the heterologous expression of the *mscS* from '*Ca.* N. maritimus' SCM1 in *E. coli*, the codon-optimized synthetic *mscS* gene was amplified from the plasmid obtained from the supplier (LifeTechnologies) by PCR using custom synthesized DNA primers. Short DNA fragments carrying NcoI and HindIII restriction sites, respectively, were attached either to the 5′ or to the 3′ prime ends of the PCR product, thereby enabling its directional insertion into the expression vector pTrc99a (Amann *et al.*, 1988). This positioned the transcription of the *mscS* gene under the control of the *lac* promoter carried by the pTrc99a vector and resulted in the isolation of plasmid pLC18. The correct nucleotide sequence of the *mscS* gene carried by pLC18 was ascertained by DNA sequence analysis, which was performed by Eurofins MWG Operon (Ebersberg, Germany).

Bacterial strains

The *E. coli* strain MJF641 which is defective in all currently known MscL-, MscS- and MscM-type channels (*mscL mscS mscK ybdG ybiO yjeP ynal*) (also referred to as Δ7) and its parent Frag1 (*thi rha lac gal*) have been described (Edwards *et al.*, 2012) and were a kind gift of Dr Ian Booth and Dr Samanta Miller (University of Aberdeen, Aberdeen, Scotland; UK). The *E. coli* B strain BL21 (Dubendorff and Studier, 1991) served as the host strain for plasmids that were used for the overexpression of recombinant versions of the '*Ca.* N. maritimus' SCM1 *ectC* and *ectD* genes. The '*Ca.* N.

maritimus' strain SCM1 (Könneke *et al.*, 2005) was from stocks of the laboratory of Dr M. Könneke (MARUM, University of Bremen; Germany).

Media and growth conditions

'*Candidatus* Nitrosopumilus maritimus' SCM1 was cultivated at 28°C in 15-l batch cultures using a 4-(2-hydroxyethyl)-1-piperazineethanesulfonate (HEPES)-buffered medium as described previously (Könneke *et al.*, 2005; Martens-Habbena *et al.*, 2009). Ammonia (1.5 mM NH_4Cl) served as the energy source, and 1 mM sodium bicarbonate was added as the carbon source. For optimal growth conditions, cells of *Ca.* N. maritimus from two batches (a total of 30 l) were grown in media with 26 g l^{-1} NaCl. This medium has an osmolarity of 928 mOsm. For growth conditions with increased salinity, cells were grown in three batches (a total of 45 l) of media containing 48 g l^{-1} NaCl. This medium has a measured osmolarity of 1598 mOsm. During incubation, cultures were slightly shaken by hand once a day to provide oxygen. A spectrophotometric assay was used to follow the growth of '*Ca.* N. maritimus' SCM1 by measuring the formation of nitrite (Strickland and Parson, 1972). Cells were harvested in the late growth phase (production of 1.1-1.3 mM nitrite) with a cross-flow filtration system equipped with a 0.1 μm pore size filter cassette (Sartocon-Slice Microsart, Sartorius, Göttingen, Germany). For ectoine and hydroxyectoine analysis, concentrated cell suspensions were fixed with 4% formaldehyde (16% formaldehyde solution (w/v) Thermo Scientific, Rockford, USA), and stored refrigerated.

For studying the expression of the *ectA* and *mscS* in '*Ca.* N. maritimus' SCM1 at different salt conditions of the growth medium, cells were cultured in 5 L batch cultures at 28°C in HEPES buffered medium (pH 7.6) containing 1.5 mM ammonia as sole energy source as described previously (Könneke *et al.*, 2014). Medium for low, optimal and high salt conditions contained 13, 26 and 39 g NaCl L^{-1} respectively. Cultures grown at different salinities were harvested in the same growth phase (when about two thirds of the ammonia was converted to nitrite) with a cross-flow filtration system (0.1 μm Sartocon slice Microsart filter, Sartorius Stedim, Göttingen, Germany). Concentrated cells were centrifuged (45 min, 4415 x *g*), and pellets were stored with RNAlater (SIGMA Life Science, Taufkirchen, Germany) at 4°C following the instructions of the manufacturer. Salt shock conditions were created by addition of 23 g NaCl L^{-1} into a 15 L batch culture of '*Ca.* N. maritimus' SCM1 in mid-growth phase grown at optimal salt conditions (26 g NaCl L^{-1}). After an additional incubation period of 24 h at high-salinity growth conditions, the cells were harvested as described above. Growth in all cultivation experiments of '*Ca.* N. maritimus' SCM1 was monitored by following the formation of nitrite. Purity of the cultures was routinely checked by phase contrast microscopy.

Escherichia coli strains were routinely maintained on LB agar plates (Miller, 1972). When strains contained plasmids, ampicillin (100 μg ml^{-1}) was added to the growth medium to select for the presence of the plasmids. For the analysis of cell viability of strains Frag1(pTrc99a), MJF641(pTrc99a) and MJF641(pLC18) subsequent to an osmotic down-shock, we used the growth medium and the procedure described by Levina and colleagues (1999). Osmotically unstressed cells were grown at 37°C in a 100-ml Erlenmeyer flask filled with 20 ml of medium in a shaking water bath (set to 220 rpm) using a citrate-phosphate buffered chemically defined medium (pH 7.0). It contained per litre: 8.58 g Na_2HPO_4, 0.87 g K_2HPO_4, 1.34 g citric acid, 1.0 g (NH_4SO_4), 0.001 g thiamine, 0.1 g $MgSO_4{\cdot}7H_2O$ and 0.002 g $(NH_4)_2SO_4{\cdot}FeSO_4{\cdot}6H_2O$, and was supplemented with 0.2% (w/v) glucose as the carbon source (Levina *et al.*, 1999). This medium possesses a measured osmolarity of 235 mOsm. For cells that were grown at high salinity, 0.3 M NaCl were added to the basal medium; it had a measured osmolarity of 730 mOsm. The osmolarity of growth media was determined with an osmometer (Vapor Pressure Osmometer 5500, Wesco, USA).

Functional complementation of the mechano-channel-defective E. coli *mutant MJF641 by the* '*Ca. N. maritimus' SCM1* MscS *protein*

The *E. coli* strains Frag1 and MJF641 (Edwards *et al.*, 2012) harbouring different plasmids were inoculated in 5 ml LB medium containing ampicillin (100 μg ml^{-1}) and were grown for 5 h. Cells were subsequently transferred into the above-described citrate-phosphate medium and incubated overnight at 37°C. Cells were then diluted to an OD_{578} of 0.05 into 20 ml of the above-described minimal medium, or into 20 ml medium that contained 0.3 M NaCl, and the cultures were subsequently grown to an OD_{578} of 0.15. At this point, IPTG was added to the cultures (final concentration 1 mM) to induce the activity of the *lac* promoter present on the backbone of the expression plasmid pTrc99a (Amann *et al.*, 1988), thereby triggering the transcription of the codon-optimized '*Ca.* N. maritimus' SCM1 *mscS* gene. Growth of the cells was subsequently continued until they reached an OD_{578} of about 0.35. These cultures were then diluted 20-fold into pre-warmed medium (37°C) containing 1 mM IPTG with (no osmotic down-shock) or without (osmotic down-shock) 0.3 M NaCl; the cells were subsequently incubated at 37°C in a shaking water bath for 30 min. To determine the number of the cells that survived the osmotic down-shock, 50-μl samples were taken after 30 min of incubation from these cultures, serial diluted in four independent sets in media with corresponding osmolarities and four 5 μl samples of the osmotically downshifted cells were then spotted onto Luria–Bertani (LB) agar plates (Miller, 1972). Those from the high osmolarity grown cells were spotted onto LB agar plates with a total NaCl content of 0.3 M NaCl. Colony-forming units were determined after overnight incubation of the LB plates at 37°C.

Overproduction, purification and determination of the quaternary assembly of the ectoine synthase and of the ectoine hydroxylase

Cells of strain BL21, harbouring either the *ectC* expression plasmid pWN48 or the *ectD* expression plasmid pMP45 were grown in a minimal medium, and the overproduction of the (*Nm*)EctC-*Strep*-tag II and (*Nm*)EctD-*Strep*-tag II proteins was initiated by adding the inducer (anhydrotetracycline) of

the TetR repressor controlled *tet* promoter to the growth medium. Cleared cell extracts of the protein overproducing cultures were prepared and used to purify the (*Nm*)EctC-*Strep*-tag II and (*Nm*)EctD-*Strep*-tag II proteins by affinity chromatography on streptactin affinity resin as detailed elsewhere (Hoeppner *et al.*, 2014; Kobus *et al.*, 2015). The purity of the recombinant (*Nm*)EctC-*Strep*-tag II and (*Nm*)EctD-*Strep*-tag II proteins was assessed by SDS-polyacrylamide gel electrophoresis (12.5% and 15% respectively); the electrophoretically separated proteins were stained with Coomassie Brilliant Blue. To analyse the quaternary assembly of the ectoine synthase and ectoine hydroxylase, we performed size-exclusion chromatography. For the analysis of the ectoine synthase, a HiLoad 16/600 Superdex 75 pg column (GE Healthcare, München, Germany) was used, and the column was run in a buffer containing 20 mM TES (pH 7.0) and 150 mM NaCl. For the analysis of the ectoine hydroxylase, a HiLoad 16/600 Superdex 200 pg column (GE Healthcare, München, Germany) was used, and the column was run in a buffer containing 20 mM TES (pH 7.5) and 200 mM KCl.

Enzyme assays of the 'Ca. N. maritimus' SCM1 ectoine synthase and of the ectoine hydroxylase

High-performance liquid chromatography-based enzyme assays (Bursy *et al.*, 2007; Widderich *et al.*, 2014a) were used to assess the biochemical and kinetic properties of the affinity-purified (*Nm*)EctC-*Strep*-tag II and (*Nm*)EctD-*Strep*-tag II proteins. Single parameters (e.g. the salt concentrations) were changed to determine optimal conditions, and variations of the substrate concentration were used to assess the kinetic parameters of the EctC. The optimized assay buffer contained 20 mM TES (pH 7.0), 150 mM NaCl, 1 mM $FeSO_4$ and 10 mM Nγ-ADABA and was run for 15 min at 30°C. Similarly, the optimized assay buffer for the EctD contained 20 mM TES (pH 7.5), 200 mM KCl, 1 mM $FeSO_4$, 10 mM 2-oxoglutarate, and 6 mM ectoine and was run for 15 min at 35°C under vigorous shaking to assure aeration, since EctD is an oxygen-dependent enzyme. To determine the kinetic parameters of the (*Nm*)EctC and (*Nm*)EctD enzymes, the substrate concentration was varied for the EctC enzyme between 0 and 50 mM; for EctD, the substrate concentration of ectoine varied between 0 mM and 50 mM while keeping the concentration of the co-substrate 2-oxoglutarate constant at 10 mM. To determine the kinetic parameters for the co-substrate of EctD, the concentration was varied between 0 mM and 40 mM, while the concentration of the ectoine substrate was kept constant at 6 mM. Ectoine synthase and EctD enzyme assays were terminated by adding acetonitrile to a final concentration of 50% to the total enzyme assay solution. Denatured proteins were removed by centrifugation (5 min) in a table-top Eppendorf centrifuge, and portions of the supernatant were then used for HPLC analysis of the formed ectoine or 5-hydroxyectoine respectively.

HPLC analysis of ectoine and hydroxyectoine content

Ectoine and hydroxyectoine were detected by HPLC analysis using an Agilent 1260 Infinity LC system (Agilent, Waldbronn, Germany), a GROM-SIL Amino 1PR column (GROM, Rottenburg-Hailfingen, Germany) essentially as described (Kuhlmann and Bremer, 2002), with the exception that a 1260 Infinity Diode Array Detector (DAD) (Agilent) was employed, instead of the previously used UV/Vis detection system. The ectoine and hydroxyectoine content of samples was quantified using the OpenLAB software suite (Agilent). When samples of EctC or EctD enzyme activity assays were measured, 5 µl to 10 µl samples were injected into the system. To determine the ectoine and hydroxyectoine content of '*Ca.* N. maritimus' SCM1 cells, samples were lysed by re-suspending them in 20% ethanol for 1 h. Cellular debris was then removed by centrifugation (13 000 r.p.m. in a Sorval centrifuge at 4°C) for 20 min, and the supernatant was subsequently lyophylized to complete dryness. The resin was re-suspended in a mixture of water : acetonitrile (50:50 v/v), and samples were then injected into the HPLC system. Since cells of '*Ca.* N. maritimus' SCM1 grown at different salinities exhibit different cell sizes, the total protein content of the originally ethanolic extract were used for a protein assay (Pierce Protein Assay Kit; ThermoScientific, Schwerte, Germany), and these values were used to standardize the ectoine/hydroxyectoine content of the cell samples.

Mapping of the transcriptional organization of the ectoine/hydroxyectoine gene cluster

To assess the transcriptional organization of the *ect* operon, cells of '*Ca.* N. maritimus' SCM1 grown under enhanced osmotic stress conditions (the medium had an osmolarity of 1598 mOsm) were used to isolate total RNA. Cell lysis was achieved by re-suspension in 20% ethanol, and total RNA was isolated from these cell extracts by using the High Pure RNA Isolation Kit (Roche, Mannheim, Germany) according to the instructions of the user manual. Samples of RNA were further purified using the RNeasy Kit (Qiagen, Hilden, Germany) as described in the user manual and used for one-step RT-PCR assays. To analyse whether the *hypectABCD-mscS* genes are transcribed in a unit, four intergenic regions of the putative operon were amplified from isolated RNA using the Qiagen One Step RT-PCR Kit and custom synthesized DNA primers (MWG, Ebersberg, Germany). As controls, we also amplified DNA regions between genes that were divergently transcribed (Fig. 7A). To ensure that the formed PCR products did not result from DNA contaminations of RNA sample used for the RT-PCR reaction, an assay was performed in which total RNA was added after the reverse transcription step.

Assessment of the transcript levels of the 'Ca. N. maritimus' SCM1 ectA and mscS genes in response to varying osmolarities

For studying the expression of the ectoine biosynthesis gene cluster in *'Ca.* N. maritimus' SCM1 under different osmotic growth conditions, cells that had been cultured either at low (13 g NaCl L^{-1}; 220 mM), optimal (26 g NaCl L^{-1}; 440 mM) or high salt (39 g NaCl L^{-1}; 660 mM) were used. Total RNA was extracted from these cells using the peqGOLD TriFast Kit (VWR International GmbH, Erlangen, Germany) according to

the manufacturer's instructions. The extracted RNA solutions were treated with RNAse-free DNAse I (Life Technologies GmbH, Darmstadt, Germany) to remove residual chromosomal DNA, again following the manufacturer's instructions. The absence of DNA contamination was ascertained by PCR analysis. The relative abundance of the *hyp-ectABCD-mscS* mRNA to the mRNA of *amoA* (Nmar_1500) was determined by real-time PCR in a CFX96 Touch Real-Time PCR Detection System (Bio-Rad Laboratories GmbH, München, Germany) with the LightCycler RNA Master SYBR green I kit (Roche Diagnostics, Mannheim, Germany). Each reaction of the one-step RT-PCR was conducted in a 20-µl volume containing 125 ng template RNA, 0.5 µM of each primer, 3.25 mM $Mn(OAc)_2$ and 7.5 µl of LightCycler RNA Master SYBR green I. The following PCR primer sets were used: ectA_fwd (5′-TTAGAGAGCCTCGAGTTGATGACGC-3′) and ectA_rev (5′-GTCAAGAGGCTTGTTGTTTTGCACC-3′), mscS_fwd (5′-CGCAAAGGAACTATTCTCAAGCTGG-3′) and mscS_rev (5′-GCGAGAAATTGAAACAAGAACCTCG-3′), and amoA_fwd (5′-CCAAGTAGGTAAGTTCTATAA-3′) and amoA_rev (5′-AAGCGGCCATCCATCTGTA-3′) as described previously (Nakagawa and Stahl, 2013). The PCR cycling conditions were used as described in the manufacturer's instructions with denaturation at 95°C for 5 s, annealing at 60°C for 10 s and extension at 72°C for 5 s. The relative expression of each gene under the tested conditions was determined by using the *amoA* transcript level in *'Ca. N. maritimus' SCM1* cells grown under optimal salt concentrations as the standard. The level of the *amoA* transcript was set to one and those of *ectA* and *mscS* are expressed in relation to this reference transcript.

Database searches and phylogenetic analysis of EctC and EctD-type proteins

The amino acid sequence of the '*Ca.* N. maritimus' SCM1 EctC protein (Walker *et al.*, 2010) was used as the template for BLAST searches of the microbial database of the JGI of the US Department of Energy (http://jgi.doe.gov/) (Nordberg *et al.*, 2013). EctC-type amino acid sequences of closely related strains of the same species were removed, and the remaining 723 retrieved amino acid sequences were aligned using the MAFFT multiple amino acid sequence alignment server (http://mafft.cbrc.jp/alignment/server/) (Katoh and Standley, 2013). This curated data set was then used to construct a rooted phylogenetic tree of EctC-type sequences by employing the iTOL software suit (http://itol.embl.de/) (Letunic and Bork, 2011).

Modelling of three-dimensional protein structures and preparation of figures of crystal structures

The amino acid sequence of the '*Ca.* N. maritimus' SCM1 MscS protein was submitted via the website of the SWISS-Model server (http://swissmodel.expasy.org/) (Biasini *et al.*, 2014) to generate an *in silico* model of the three-dimensional structure of this protein. The program automatically chose the monomer of the heptameric *E. coli* MscS open crystal structure (Wang *et al.*, 2008) as the modelling template (PDB accession code: 2VV5). The model of the monomer of the (*Nm*)MscS protein generated via the SWISS-Model web server was visualized and compared with the *E. coli* MscS protein using resources provided by the PyMol suit (https://www.pymol.org/).

Acknowledgements

We thank Irina Bagyan for the kind gift of ectoines and Ian Booth and Samata Miller for generously providing bacterial strains. We greatly appreciate the expert technical assistance of Jochen Sohn and Jutta Gade in the course of this project and highly value the kind help of Vickie Koogle in the language editing of our manuscript. We thank Tamara Hoffmann for helpful discussions and critical reading of the manuscript. Funding for this study was provided by the Deutsche Forschungsgemeinschaft (DFG) through the SFB 987 (to E.B. and J.H.) (Marburg), the Transregio SFB 51 (to J.S.D.) (Braunschweig/Oldenburg), The G.W. Leibniz Prize (to K.-U. Hinrichs; MARUM, Bremen) and through the LOEWE Program of the State of Hessen (via the Centre for Synthetic Microbiology; Synmicro; Marburg) (to J.H. and E.B.). N.W. and L.C. are recipients of PhD fellowships from the International Max Planck Research School for Environmental, Cellular and Molecular Microbiology (IMPRS-Mic, Marburg) and gratefully acknowledge its generous support.

References

Amann, E., Ochs, B., and Abel, K.J. (1988) Tightly regulated tac promoter vectors useful for the expression of unfused and fused proteins in *Escherichia coli*. *Gene* **69:** 301–315.

Bass, R.B., Strop, P., Barclay, M., and Rees, D.C. (2002) Crystal structure of *Escherichia coli* MscS, a voltage-modulated and mechanosensitive channel. *Science* **298:** 1582–1587.

Bayer, B., Vojvoda, J., Offre, P., Alves, R.J., Elisabeth, N.H., Garcia, J.A., *et al.* (2015) Physiological and genomic characterization of two novel marine thaumarchaeal strains indicates niche differentiation. *ISME J* in press. doi: 10.1038/ismej.2015.200

Becker, E.A., Seitzer, P.M., Tritt, A., Larsen, D., Krusor, M., Yao, A.I., *et al.* (2014) Phylogenetically driven sequencing of extremely halophilic archaea reveals strategies for static and dynamic osmo-response. *PLoS Genet* **10:** e1004784.

Becker, M., and Krämer, R. (2015) MscCG from *Corynebacterium glutamicum*: functional significance of the C-terminal domain. *Eur Biophys J* **44:** 577–588.

Becker, M., Borngen, K., Nomura, T., Battle, A.R., Marin, K., Martinac, B., and Krämer, R. (2013) Glutamate efflux mediated by *Corynebacterium glutamicum* MscCG, *Escherichia coli* MscS, and their derivatives. *Biochim Biophys Acta* **1828:** 1230–1240.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., *et al.* (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* **42:** W252–W258.

Blainey, P.C., Mosier, A.C., Potanina, A., Francis, C.A., and Quake, S.R. (2011) Genome of a low-salinity ammonia-oxidizing archaeon determined by single-cell and metagenomic analysis. *PLoS ONE* **6:** e16626.

Bolen, D.W., and Baskakov, I.V. (2001) The osmophobic effect: natural selection of a thermodynamic force in protein folding. *J Mol Biol* **310:** 955–963.

Booth, I.R. (2014) Bacterial mechanosensitive channels: progress towards an understanding of their roles in cell physiology. *Curr Opin Microbiol* **18:** 16–22.

Booth, I.R., and Blount, P. (2012) The MscS and MscL families of mechanosensitive channels act as microbial emergency release valves. *J Bacteriol* **194:** 4802–4809.

Booth, I.R., Miller, S., Müller, A., and Lehtovirta-Morley, L. (2015) The evolution of bacterial mechanosensitive channels. *Cell Calcium* **57:** 140–150.

Börngen, K., Battle, A.R., Möker, N., Morbach, S., Marin, K., Martinac, B., and Krämer, R. (2010) The properties and contribution of the *Corynebacterium glutamicum* MscS variant to fine-tuning of osmotic adaptation. *Biochim Biophys Acta* **1798:** 2141–2149.

Bremer, E., and Krämer, R. (2000) Coping with osmotic challenges: osmoregulation through accumulation and release of compatible solutes. In *Bacterial Stress Responses*. Storz, G., and Hengge-Aronis, R. (eds). Washington DC, USA: ASM Press, pp. 79–97.

Brill, J., Hoffmann, T., Bleisteiner, M., and Bremer, E. (2011) Osmotically controlled synthesis of the compatible solute proline is critical for cellular defense of *Bacillus subtilis* against high osmolarity. *J Bacteriol* **193:** 5335–5346.

Brochier-Armanet, C., Boussau, B., Gribaldo, S., and Forterre, P. (2008) Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat Rev Microbiol* **6:** 245–252.

Brown, A.D. (1976) Microbial water stress. *Bacteriol Rev* **40:** 803–846.

Bursy, J., Pierik, A.J., Pica, N., and Bremer, E. (2007) Osmotically induced synthesis of the compatible solute hydroxyectoine is mediated by an evolutionarily conserved ectoine hydroxylase. *J Biol Chem* **282:** 31147–31155.

Bursy, J., Kuhlmann, A.U., Pittelkow, M., Hartmann, H., Jebbar, M., Pierik, A.J., and Bremer, E. (2008) Synthesis and uptake of the compatible solutes ectoine and 5-hydroxyectoine by *Streptomyces coelicolor* A3(2) in response to salt and heat stresses. *Appl Environ Microbiol* **74:** 7286–7296.

Calderon, M.I., Vargas, C., Rojo, F., Iglesias-Guerra, F., Csonka, L.N., Ventosa, A., and Nieto, J.J. (2004) Complex regulation of the synthesis of the compatible solute ectoine in the halophilic bacterium *Chromohalobacter salexigens* DSM 3043^{T}. *Microbiology* **150:** 3051–3063.

Canovas, D., Vargas, C., Iglesias-Guerra, F., Csonka, L.N., Rhodes, D., Ventosa, A., and Nieto, J.J. (1997) Isolation and characterization of salt-sensitive mutants of the moderate halophile *Halomonas elongata* and cloning of the ectoine synthesis genes. *J Biol Chem* **272:** 25794–25801.

Coquelle, N., Talon, R., Juers, D.H., Girard, E., Kahn, R., and Madern, D. (2010) Gradual adaptive changes of a protein facing high salt concentrations. *J Mol Biol* **404:** 493–505.

da Costa, M.S., Santos, H., and Galinski, E.A. (1998) An overview of the role and diversity of compatible solutes in Bacteria and Archaea. *Adv Biochem Eng Biotechnol* **61:** 117–153.

Cruickshank, C.C., Minchin, R.F., Le Dain, A.C., and Martinac, B. (1997) Estimation of the pore size of the large-conductance mechanosensitive ion channel of *Escherichia coli*. *Biophys J* **73:** 1925–1931.

Csonka, L.N. (1989) Physiological and genetic responses of bacteria to osmotic stress. *Microbiol Rev* **53:** 121–147.

Deole, R., Challacombe, J., Raiford, D.W., and Hoff, W.D. (2013) An extremely halophilic proteobacterium combines a highly acidic proteome with a low cytoplasmic potassium content. *J Biol Chem* **288:** 581–588.

Deschamps, P., Zivanovic, Y., Moreira, D., Rodriguez-Valera, F., and Lopez-Garcia, P. (2014) Pangenome evidence for extensive interdomain horizontal transfer affecting lineage core and shell genes in uncultured planktonic thaumarchaeota and euryarchaeota. *Genome Biol Evol* **6:** 1549–1563.

Dong, C., Major, L.L., Srikannathasan, V., Errey, J.C., Giraud, M.F., Lam, J.S., *et al.* (2007) RmlC, a C3′ and C5′ carbohydrate epimerase, appears to operate via an intermediate with an unusual twist boat conformation. *J Mol Biol* **365:** 146–159.

Dubendorff, J.W., and Studier, F.W. (1991) Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with lac repressor. *J Mol Biol* **219:** 45–59.

Dunwell, J.M., Culham, A., Carter, C.E., Sosa-Aguirre, C.R., and Goodenough, P.W. (2001) Evolution of functional diversity in the cupin superfamily. *Trends Biochem Sci* **26:** 740–746.

Edwards, M.D., Black, S., Rasmussen, T., Rasmussen, A., Stokes, N.R., Stephen, T.L., *et al.* (2012) Characterization of three novel mechanosensitive channel activities in *Escherichia coli*. *Channels* **6:** 272–281.

Elling, F.J., Könneke, M., Lipp, J.S., Becker, K.W., Gagen, E.J., and Hinrichs, K.-U. (2014) Effect of growth phase on the membrane lipid composition of the thaumarchaeon *Nitrosopumilus martitimus* and their implication for archaeal lipid distribution in the marine environment. *Geochim Cosmochim Acta* **141:** 579–597.

Elling, F.J., Könneke, M., Mußmann, M., Greve, A., and Hinrichs, K.-U. (2015) Influence of temperature, pH, and salinity on membrane lipid composition and Tex86 of marine planktonic thaumarchaealisolates. *Geochim Cosmochim Acta* in press. doi:10.1016/j.gca.2015.09.004

Empadinhas, N., and da Costa, M.S. (2011) Diversity, biological roles and biosynthetic pathways for sugar-glycerate containing compatible solutes in bacteria and archaea. *Environ Microbiol* **13:** 2056–2077.

Galinski, E.A., and Trüper, H.G. (1994) Microbial behaviour in salt-stressed ecosystems. *FEMS Microbiol Rev* **15:** 95–108.

Galinski, E.A., Pfeiffer, H.P., and Trüper, H.G. (1985) 1,4,5,6-Tetrahydro-2-methyl-4-pyrimidinecarboxylic acid. A novel cyclic amino acid from halophilic phototrophic bacteria of the genus *Ectothiorhodospira*. *Eur J Biochem* **149:** 135–139.

Garcia-Estepa, R., Argandona, M., Reina-Bueno, M., Capote, N., Iglesias-Guerra, F., Nieto, J.J., and Vargas, C. (2006) The *ectD* gene, which is involved in the synthesis of the compatible solute hydroxyectoine, is essential for thermoprotection of the halophilic bacterium *Chromohalobacter salexigens*. *J Bacteriol* **188:** 3774–3784.

Gubry-Rangin, C., Kratsch, C., Williams, T.A., McHardy, A.C., Embley, T.M., Prosser, J.I., and Macqueen, D.J. (2015)

Coupling of diversification and pH adaptation during the evolution of terrestrial Thaumarchaeota. *Proc Natl Acad Sci USA* **112:** 9370–9375.

Harishchandra, R.K., Wulff, S., Lentzen, G., Neuhaus, T., and Galla, H.J. (2010) The effect of compatible solute ectoines on the structural organization of lipid monolayer and bilayer membranes. *Biophys Chem* **150:** 37–46.

Haswell, E.S., Phillips, R., and Rees, D.C. (2011) Mechanosensitive channels: What can they do and how do they do it? *Structure* **19:** 1356–1369.

Hoeppner, A., Widderich, N., Bremer, E., and Smits, S.H.J. (2014) Overexpression, crystallization and preliminary X-ray crystallographic analysis of the ectoine hydroxylase from *Sphingopyxis alaskensis*. *Acta Cryst* **F70:** 493–496.

Hoffmann, T., Wensing, A., Brosius, M., Steil, L., Völker, U., and Bremer, E. (2013) Osmotic control of *opuA* expression in *Bacillus subtilis* and its modulation in response to intracellular glycine betaine and proline pools. *J Bacteriol* **195:** 510–522.

Höppner, A., Widderich, N., Lenders, M., Bremer, E., and Smits, S.H.J. (2014) Crystal structure of the ectoine hydroxylase, a snapshot of the active site. *J Biol Chem* **289:** 29570–29583.

Inbar, L., and Lapidot, A. (1988) The structure and biosynthesis of new tetrahydropyrimidine derivatives in actinomycin D producer *Streptomyces parvulus*. Use of ^{13}C- and ^{15}N-labeled L-glutamate and ^{13}C and ^{15}N NMR spectroscopy. *J Biol Chem* **263:** 16014–16022.

Katoh, K., and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30:** 772–780.

Kelly, W.J., Leahy, S.C., Li, D., Perry, R., Lambie, S.C., Attwood, G.T., and Altermann, E. (2014) The complete genome sequence of the rumen methanogen *Methanobacterium formicicum* BRM9. *Stand Genomic Sci* **9:** 15.

Kempf, B., and Bremer, E. (1998) Uptake and synthesis of compatible solutes as microbial stress responses to high osmolarity environments. *Arch Microbiol* **170:** 319–330.

Klähn, S., and Hagemann, M. (2011) Compatible solute biosynthesis in cyanobacteria. *Environ Microbiol* **13:** 551–562.

Kloda, A., and Martinac, B. (2002) Mechanosensitive channels of bacteria and archaea share a common ancestral origin. *Eur Biophys J* **31:** 14–25.

Kobus, S., Widderich, N., Hoeppner, A., Bremer, E., and Smits, S.H.J. (2015) Overproduction, crystallization and X-ray diffraction data analysis of ectoine synthase from the cold-adapted marine bacterium *Sphingopyxis alaskensis*. *Acta Cryst* **F71:** 1027–1032.

Kol, S., Merlo, M.E., Scheltema, R.A., de Vries, M., Vonk, R.J., Kikkert, N.A., *et al.* (2010) Metabolomic characterization of the salt stress response in *Streptomyces coelicolor*. *Appl Environ Microbiol* **76:** 2574–2581.

Könneke, M., Bernhard, A.E., de la Torre, J.R., Walker, C.B., Waterbury, J.B., and Stahl, D.A. (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437:** 543–546.

Könneke, M., Schubert, D.M., Brown, P.C., Hugler, M., Standfest, S., Schwander, T., *et al.* (2014) Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO2 fixation. *Proc Natl Acad Sci USA* **111:** 8239–8244.

Kuhlmann, A.U., and Bremer, E. (2002) Osmotically regulated synthesis of the compatible solute ectoine in *Bacillus pasteurii* and related Bacillus spp. *Appl Environ Microbiol* **68:** 772–783.

Kuhlmann, A.U., Bursy, J., Gimpel, S., Hoffmann, T., and Bremer, E. (2008) Synthesis of the compatible solute ectoine in *Virgibacillus pantothenticus* is triggered by high salinity and low growth temperature. *Appl Environ Microbiol* **74:** 4560–4563.

Kunte, H.J., Galinski, E.A., and Trüper, G.H. (1993) A modified FMOC-method for the detection of amino acid-type osmolytes and tetrahydropyrimidines (ectoines). *J Microbiol Meth* **17:** 129–136.

Kunte, H.J., Lentzen, G., and Galinski, E. (2014) Industrial production of the cell protectant ectoine: protection, mechanisms, processes, and products. *Curr Biotechnol* **3:** 10–25.

Kurz, M., Burch, A.Y., Seip, B., Lindow, S.E., and Gross, H. (2010) Genome-driven investigation of compatible solute biosynthesis pathways of *Pseudomonas syringae* pv. syringae and their contribution to water stress tolerance. *Appl Environ Microbiol* **76:** 5452–5462.

Letunic, I., and Bork, P. (2011) Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res* **39:** W475–W478.

Levina, N., Totemeyer, S., Stokes, N.R., Louis, P., Jones, M.A., and Booth, I.R. (1999) Protection of *Escherichia coli* cells against extreme turgor by activation of MscS and MscL mechanosensitive channels: identification of genes required for MscS activity. *EMBO J* **18:** 1730–1737.

Lippert, K., and Galinski, E.A. (1992) Enzyme stabilization by ectoine-type compatible solutes: protection against heating, freezing and drying. *Appl Microbiol Biotechnol* **37:** 61–65.

Lo, C.C., Bonner, C.A., Xie, G., D'Souza, M., and Jensen, R.A. (2009) Cohesion group approach for evolutionary analysis of aspartokinase, an enzyme that feeds a branched network of many biochemical pathways. *Microbiol Mol Biol Rev* **73:** 594–651.

Louis, P., and Galinski, E.A. (1997) Characterization of genes for the biosynthesis of the compatible solute ectoine from *Marinococcus halophilus* and osmoregulated expression in *Escherichia coli*. *Microbiology* **143:** 1141–1149.

Manzanera, M., Garcia de Castro, A., Tondervik, A., Rayner-Brandes, M., Strom, A.R., and Tunnacliffe, A. (2002) Hydroxyectoine is superior to trehalose for anhydrobiotic engineering of *Pseudomonas putida* KT2440. *Appl Environ Microbiol* **68:** 4328–4333.

Martens-Habbena, W., Berube, P.M., Urakawa, H., de la Torre, J.R., and Stahl, D.A. (2009) Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature* **461:** 976–979.

Miller, J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor New York: Cold Spring Harbor Laboratory.

Moritz, K.D., Amendt, B., Witt, E.M.H.J., and Galinski, E.A. (2015) The hydroxyectoine gene cluster of the non-halophilic acidophile *Acidiphilium cryptum*. *Extremophiles* **19:** 87–99.

Müller, V., Spanheimer, R., and Santos, H. (2005) Stress response by solute accumulation in archaea. *Curr Opin Microbiol* **8:** 729–736.

Naismith, J.H., and Booth, I.R. (2012) Bacterial mechanosensitive channels – MscS: Evolution's solution to creating sensitivity in function. *Annu Rev Biophys* **41:** 157–177.

Nakagawa, T., and Stahl, D.A. (2013) Transcriptional response of the archaeal ammonia oxidizer *Nitrosopumilus maritimus* to low and environmentally relevant ammonia concentrations. *Appl Environ Microbiol* **79:** 6911–6916.

Nakamura, J., Hirano, S., Ito, H., and Wachi, M. (2007) Mutations of the *Corynebacterium glutamicum* NCgl1221 gene, encoding a mechanosensitive channel homolog, induce L-glutamic acid production. *Appl Environ Microbiol* **73:** 4491–4498.

Nakayama, Y., Yoshimura, K., and Iida, H. (2013) Electrophysiological characterization of the mechanosensitive channel MscCG in *Corynebacterium glutamicum*. *Biophys J* **105:** 1366–1375.

Nelson-Sathi, S., Sousa, F.L., Roettger, M., Lozada-Chavez, N., Thiergart, T., Janssen, A., *et al.* (2015) Origins of major archaeal clades correspond to gene acquisitions from bacteria. *Nature* **517:** 77–80.

Nordberg, H., Cantor, M., Dusheyko, S., Hua, S., Poliakov, A., Shabalov, I., *et al.* (2013) The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. *Nucleic Acids Res* **42:** D26–D31.

Ochman, H., Lawrence, J.G., and Groisman, E.A. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* **405:** 299–304.

Offre, P., Spang, A., and Schleper, C. (2013) Archaea in biogeochemical cycles. *Annu Rev Microbiol* **67:** 437–457.

Ono, H., Okuda, M., Tongpim, S., Imai, K., Shinmyo, A., Sakuda, S., *et al.* (1998) Accumulation of compatible solutes, ectoine and hydroxyectoine, in a moderate halophile, *Halomonas elongata* KS3, isolated from dry salty land in Thailand. *J Ferment Bioeng* **85:** 362–368.

Ono, H., Sawada, K., Khunajakr, N., Tao, T., Yamamoto, M., Hiramoto, M., *et al.* (1999a) Characterization of biosynthetic enzymes for ectoine as a compatible solute in a moderately halophilic eubacterium, *Halomonas elongata*. *J Bacteriol* **181:** 91–99.

Ono, H., Sawada, K., Khunajakr, N., Tao, T., Yamamoto, M., Hiramoto, M., *et al.* (1999b) Characterization of biosynthetic enzymes for ectoine as a compatible solute in a moderately halophilic eubacterium, *Halomonas elongata*. *J Bacteriol* **181:** 91–99.

Oren, A. (2011) Thermodynamic limits to microbial life at high salt concentrations. *Environ Microbiol* **13:** 1908–1923.

Oren, A. (2013) Life at high salt concentrations, intracellular KCl concentrations, and acidic proteomes. *Front Microbiol* **4:** 315.

Pastor, J.M., Salvador, M., Argandona, M., Bernal, V., Reina-Bueno, M., Csonka, L.N., *et al.* (2010) Ectoines in cell stress protection: uses and biotechnological production. *Biotechnol Adv* **28:** 782–801.

Pester, M., Schleper, C., and Wagner, M. (2011) The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. *Curr Opin Microbiol* **14:** 300–306.

Peters, P., Galinski, E.A., and Trüper, H.G. (1990) The biosyntheis of ectoine. *FEMS Microbiol Lett* **71:** 157–162.

Pivetti, C.D., Yen, M.R., Miller, S., Busch, W., Tseng, Y.H., Booth, I.R., and Saier, M.H., Jr (2003) Two families of mechanosensitive channel proteins. *Microbiol Mol Biol Rev* **67:** 66–85.

Prabhu, J., Schauwecker, F., Grammel, N., Keller, U., and Bernhard, M. (2004) Functional expression of the ectoine hydroxylase gene (*thpD*) from *Streptomyces chrysomallus* in *Halomonas elongata*. *Appl Environ Microbiol* **70:** 3130–3132.

Record, M.T., Jr, Courtenay, E.S., Cayley, D.S., and Guttman, H.J. (1998) Responses of *E. coli* to osmotic stress: large changes in amounts of cytoplasmic solutes and water. *Trends Biochem Sci* **23:** 143–148.

Roberts, M.F. (2004) Osmoadaptation and osmoregulation in archaea: update 2004. *Front Biosci* **9:** 1999–2019.

Roesser, M., and Müller, V. (2001) Osmoadaptation in bacteria and archaea: common principles and differences. *Environ Microbiol* **3:** 743–754.

Rowe, I., Anishkin, A., Kamaraju, K., Yoshimura, K., and Sukharev, S. (2014) The cytoplasmic cage domain of the mechanosensitive channel MscS is a sensor of macromolecular crowding. *J Gen Physiol* **143:** 543–557.

Salvador, M., Argandona, M., Pastor, J.M., Bernal, V., Canovas, M., Csonka, L.N., *et al.* (2015) Contribution of RpoS to metabolic efficiency and ectoines synthesis during the osmo- and heat-stress response in the halophilic bacterium *Chromohalobacter salexigens*. *Environ Microbiol Rep* **7:** 301–311.

Santoro, A.E., Francis, C.A., de Sieyes, N.R., and Boehm, A.B. (2008) Shifts in the relative abundance of ammonia-oxidizing bacteria and archaea across physicochemical gradients in a subterranean estuary. *Environ Microbiol* **10:** 1068–1079.

Saum, S.H., and Müller, V. (2008) Growth phase-dependent switch in osmolyte strategy in a moderate halophile: ectoine is a minor osmolyte but major stationary phase solute in *Halobacillus halophilus*. *Environ Microbiol* **10:** 716–726.

Schwibbert, K., Marin-Sanguino, A., Bagyan, I., Heidrich, G., Lentzen, G., Seitz, H., *et al.* (2011) A blueprint of ectoine metabolism from the genome of the industrial producer *Halomonas elongata* DSM 2581^{T}. *Environ Microbiol* **13:** 1973–1994.

Soucy, S.M., Huang, J., and Gogarten, J.P. (2015) Horizontal gene transfer: building the web of life. *Nat Rev Genet* **16:** 472–482.

Stahl, D.A., and de la Torre, J.R. (2012) Physiology and diversity of ammonia-oxidizing archaea. *Annu Rev Microbiol* **66:** 83–101.

Stieglmeier, M., Alves, R.J.E., and Schleper, C. (2014) The phylum Thaumarchaeota. In *The Prokaryotes – Other Major Lineages of Bacteria and the Archaea*. Rosenberg, E.I. (ed.). Berlin, Germany: Springer, pp. 347–362.

Storz, G., and Hengge-Aronis, R. (2000) *Bacterial Stress Responses*. Washington, DC, USA: ASM Press.

Stöveken, N., Pittelkow, M., Sinner, T., Jensen, R.A., Heider, J., and Bremer, E. (2011) A specialized aspartokinase enhances the biosynthesis of the osmoprotectants ectoine

and hydroxyectoine in *Pseudomonas stutzeri* A1501. *J Bacteriol* **193:** 4456–4468.
Strickland, J.D.H., and Parson, T.R. (1972) *A Practical Handbook of Seawater Analysis*. Ottawa, Canada: Fisheries Research Board of Canada.
Talon, R., Coquelle, N., Madern, D., and Girard, E. (2014) An experimental point of view on hydration/solvation in halophilic proteins. *Front Microbiol* **5:** 66.
Tanne, C., Golovina, E.A., Hoekstra, F.A., Meffert, A., and Galinski, E.A. (2014) Glass-forming property of hydroxyectoine is the cause of its superior function as a dessication protectant. *Front Microbiol* **5:** 150.
Walker, C.B., de la Torre, J.R., Klotz, M.G., Urakawa, H., Pinel, N., Arp, D.J., *et al.* (2010) *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proc Natl Acad Sci USA* **107:** 8818–8823.
Wang, W., Black, S.S., Edwards, M.D., Miller, S., Morrison, E.L., Bartlett, W., *et al.* (2008) The structure of an open form of an *E. coli* mechanosensitive channel at 3.45 A resolution. *Science* **321:** 1179–1183.
Widderich, N., Höppner, A., Pittelkow, M., Heider, J., Smits, S.H.J., and Bremer, E. (2014a) Biochemical properties of ectoine hydroxylases from extremophiles and their wider taxonomic distribution among microorganisms. *PLoS ONE* **9:** e93809.
Widderich, N., Pittelkow, M., Höppner, A., Mulnaes, D., Buckel, W., Gohlke, H., *et al.* (2014b) Molecular dynamics simulations and structure-guided mutagenesis provide insight into the architecture of the catalytic core of the ectoine hydroxylase. *J Mol Biol* **426:** 586–600.
Wilson, M.E., Maksaev, G., and Haswell, E.S. (2013) MscS-like mechanosensitive channels in plants and microbes. *Biochemistry* **52:** 5708–5722.
Wood, J.M. (2011) Bacterial osmoregulation: a paradigm for the study of cellular homeostasis. *Annu Rev Microbiol* **65:** 215–238.
Wood, J.M., Bremer, E., Csonka, L.N., Krämer, R., Poolman, B., van der Heide, T., and Smith, L.T. (2001) Osmosensing and osmoregulatory compatible solute accumulation by bacteria. *Comp Biochem Physiol A Mol Integr Physiol* **130:** 437–460.
Youssef, N.H., Savage-Ashlock, K.N., McCully, A.L., Luedtke, B., Shaw, E.I., Hoff, W.D., and Elshahed, M.S. (2014) Trehalose/2-sulfotrehalose biosynthesis and glycine-betaine uptake are widely spread mechanisms for osmoadaptation in the *Halobacteriales*. *ISME J* **8:** 636–649.
Zaparty, M., Hagemann, A., Brasen, C., Hensel, R., Lupas, A.N., Brinkmann, H., and Siebers, B. (2013) The first prokaryotic trehalose synthase complex identified in the hyperthermophilic crenarchaeon *Thermoproteus tenax*. *PLoS ONE* **8:** e61354.
Zhang, Y., Chen, L., Dai, T., Tian, J., and Wen, D. (2015) The influence of salinity on the abundance, transcriptional activity, and diversity of AOA and AOB in an estuarine sediment: a microcosm study. *Appl Microbiol Biotechnol* **99:** 9825–9833.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Fig. S1. Phylogenetic tree of EctC-type proteins. Based on an amino acid sequence alignment of 723 EctC-type proteins, a rooted phylogenetic tree was constructed with the ITOL program (Letunic and Bork, 2011). The three regions in the phylogenetic tree populated by archaeal EctC proteins are highlighted by red arrowheads. The position of those 24 EctC-type proteins that originate from bacteria that lack identifiable *ectAB* genes are marked above the phylogenetic tree by a black bar.
Fig. S2. Alignment of the amino acid sequence of the *E. coli* MscS protein with those MscS-type proteins whose corresponding genes are located next to ectoine/hydroxyectoine biosynthetic gene clusters in the genomes of '*Ca.* Nitrosopumilus maritimus' SCM 1 and '*Ca.* Nistrosopumilus' sp. AR2 (see Fig. 1B in the main text). Red bars indicate the membrane-spanning segments of the *E. coli* MscS protein.
Table S1. Amino acid sequence relatedness of the ectoine biosynthetic proteins from '*Ca.* Nitrosopumilus' strains to other ectoine biosynthetic enzymes.
Table S2. Distribution of mechanosensitive channels in potential ectoine-producing *Archaea*.

Strangers in the archaeal world: osmostress-responsive biosynthesis of ectoine and hydroxyectoine by the marine thaumarchaeon *Nitrosopumilus maritimus*

Nils Widderich[1,☯], Laura Czech[1,☯], Felix J. Elling[2], Martin Könneke[2], Nadine Stöveken[1,3], Marco Pittelkow[1], Ramona Riclea[4,5], Jeroen S. Dickschat[4,5], Johann Heider[1,3] and Erhard Bremer[1,3*]

[1]Laboratory for Molecular Microbiology, Department of Biology, Philipps-University, Karl-von-Frisch Str. 8, D-35043 Marburg, Germany

[2]Organic Geochemistry Group, MARUM – Center for Marine Environmental Sciences, University of Bremen, PO Box 330 440, D-28334 Bremen, Germany

[3]LOEWE-Center for Synthetic Microbiology, Philipps-University Marburg, Hans-Meerwein-Str. 6, D-35043 Marburg, Germany

[4]Kekulé-Institut for Organic Chemistry and Biochemistry, Friedrich-Wilhelms-University Bonn, Gerhard-Domagk Str. 1, D-53121 Bonn, Germany

[5]Institute of Organic Chemistry, TU Braunschweig, Hagenring 30, D-38106 Braunschweig, Germany

Running title: Ectoine and hydroxyectoine biosynthesis in *Archaea*

☯These authors contributed equally

*Corresponding author. Mailing address: Philipps-University Marburg, Dept. of Biology, Laboratory for Microbiology, Karl-von-Frisch-Str. 8, D-35032 Marburg, Germany. Phone: (+49)-6421-2821529. Fax: (+49)-6421-2828979. E-Mail: bremer@staff.uni-marburg.de

Supplemental Figure S1

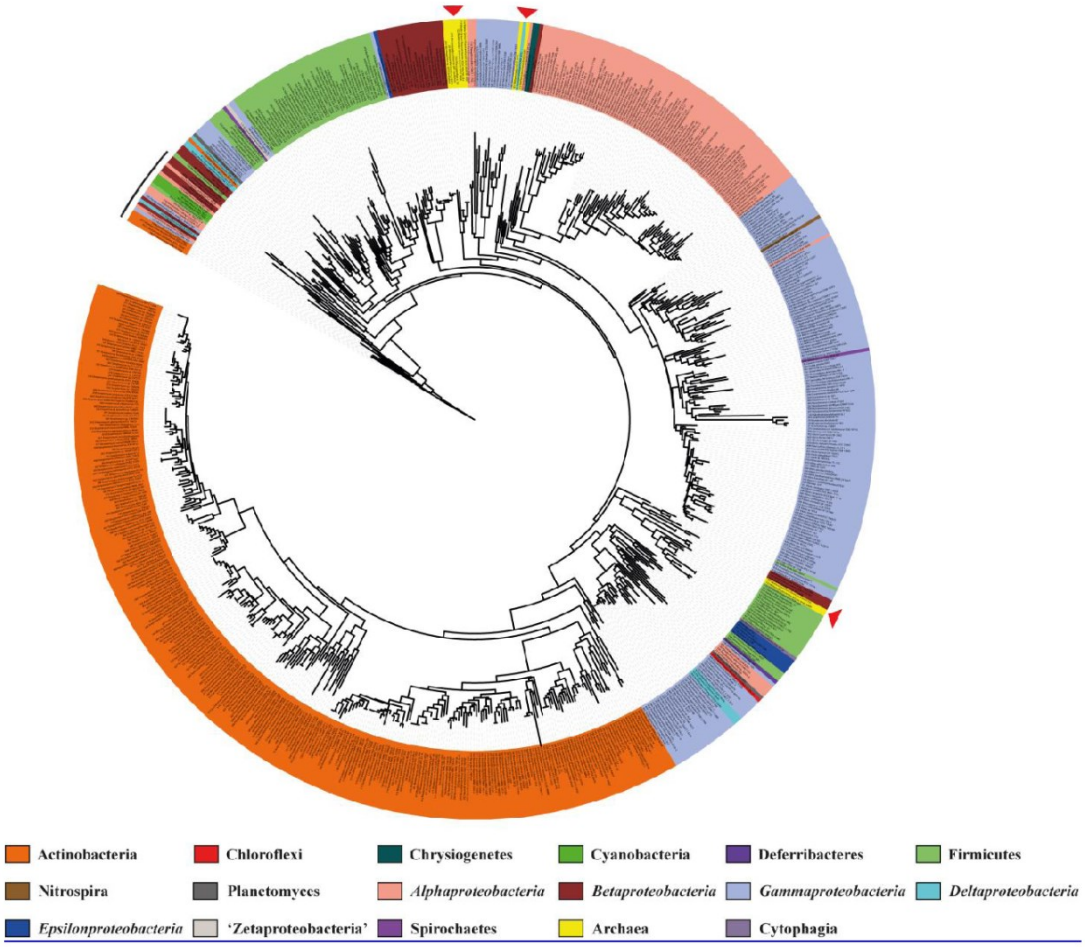


Fig. S1. Phylogenetic tree of EctC-type proteins. Based on an amino acid sequence alignment of 723 EctC-type proteins, a rooted phylogenetic tree was constructed with the iTOL program (Letunic and Bork, 2011). The three regions in the phylogenetic tree populated by archaeal EctC proteins are highlighted by red arrow-heads. The position of those 24 EctC-type proteins that originate from bacteria that lack identifiable *ectAB* genes are marked above the phylogenetic tree by a black bar.

Supplemental Figure S2

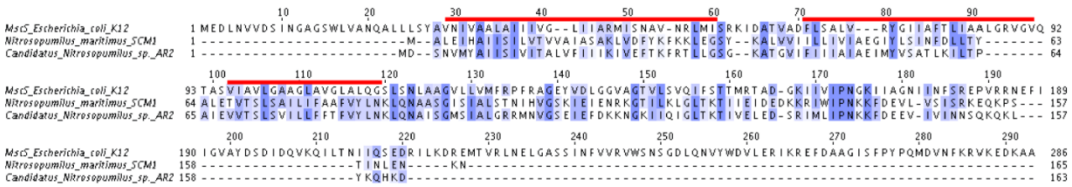


Fig. S2. Alignment of the amino acid sequence of the *E. coli* MscS protein with those MscS-type proteins whose corresponding genes are located next to ectoine/hydroxyectoine biosynthetic gene clusters in the genomes of "*Ca.* Nitrosopumilus maritimus" SCM 1 and "*Ca.* Nistrosopumilus" sp. AR2 (see Fig. 1B in the main text). Red bars indicate the membrane-spanning segments of the *E. coli* MscS protein.

Supplemental Table S1

Table S1. Amino acid sequence relatedness of the ectoine biosynthetic proteins from "*Ca*. Nitrosopumilus" strains to other ectoine biosynthetic enzymes.

EctA	"*Ca.* Nitrosopumilus maritimus" SCM1	"*Ca.* Nitrosopumilus" sp. SJ	"*Ca.* Nitrosopumilus koreensis" AR1	"*Ca.* Nitrosopumilus" sp. AR2
1	*Nitrosopumilus maritimus* (100%)	Multispecies *Nitrosopumilus* (100%)	Multispecies *Nitrosopumilus* (100%)	Multispecies *Nitrosopumilus* (100%)
2	Multispecies *Nitrosopumilus* (82%)	Marine Group I thaumarchaeote SCGC AAA799-N04 (83%)	Marine Group I thaumarchaeote SCGC AAA799-N04 (83%)	*Nitrosopumilus maritimus* (82%)
3	Marine Group I thaumarchaeote SCGC AAA799-N04 (80%)	Marine Group I thaumarchaeote SCGC AAA799-P11 (83%)	Marine Group I thaumarchaeote SCGC AAA799-P11 (83%)	Marine Group I thaumarchaeote SCGC AAA799-P11 (78%)
4	Marine Group I thaumarchaeote SCGC AAA799-P11 (80%)	Multispecies *Nitrosopumilus* (73%)	Multispecies *Nitrosopumilus* (73%)	Marine Group I thaumarchaeote SCGC AAA799-N04 (78%)
5	Multispecies *Nitrosopumilus* (74%)	*Nitrosopumilus maritimus* (74%)	*Nitrosopumilus maritimus* (74%, 6e-84)	Multispecies *Nitrosopumilus* (73%)
6	Marine Group I thaumarchaeote SCGC AAA799-E16 (79%)	Marine Group I thaumarchaeote SCGC AAA799-E16 (81%)	Marine Group I thaumarchaeote SCGC AAA799-E16 (81%)	Marine Group I thaumarchaeote SCGC AAA799-E16 (78%)
7	*Bacillus thermotolerans* (41%)	*Glycomyces arizonensis* (42%)	*Glycomyces arizonensis* (42%)	*Dehalogenimonas lykanthroporepellens* (46%)
8	*Alkalibacillus haloalkaliphilus* (44%)	*Nocardia brasiliensis* (40%)	*Nocardia brasiliensis* (40%)	*Paenibacillus daejeonensis* (41%)
9	*Oceanobacillus picturae* (47%)	*Nocardia* sp. W9851 (40%)	*Nocardia* sp. W9851 (40%)	*Methylobacter marinus* (39%)
10	*Oceanobacillus picturae* (47%)	*Marinospirillum minutulum* (43%)	*Marinospirillum minutulum* (43%)	zeta proteobacterium SCGC AB-137-C09 (42%)

EctB	"*Ca.* Nitrosopumilus maritimus" SCM1	"*Ca.* Nitrosopumilus" sp. SJ	"*Ca.* Nitrosopumilus koreensis" AR1	"*Ca.* Nitrosopumilus" sp. AR2
1	*Nitrosopumilus maritimus* (100%)	Multispecies *Nitrosopumilus* (99%)	Multispecies *Nitrosopumilus* (100%)	Multispecies *Nitrosopumilus* (100%)
2	Multispecies *Nitrosopumilus* (100%)	Marine Group I thaumarchaeote SCGC AAA799-N04 (86%)	Marine Group I thaumarchaeote SCGC AAA799-N04 (86%)	Marine Group I thaumarchaeote SCGC AAA799-N04 (80%)
3	Marine Group I thaumarchaeote SCGC AAA799-N04 (83%)	Marine Group I thaumarchaeote SCGC AAA799-P11 (86%)	Marine Group I thaumarchaeote SCGC AAA799-P11 (86%)	Marine Group I thaumarchaeote SCGC AAA799-P11 (79%)
4	Marine Group I thaumarchaeote SCGC AAA799-P11 (82%)	*Nitrosopumilus maritimus* (82%)	*Nitrosopumilus maritimus* (82%)	*Nitrosopumilus maritimus* (79%)
5	Multispecies *Nitrosopumilus* (79%)	Multispecies *Nitrosopumilus* (77%)	Multispecies *Nitrosopumilus* (78%)	Multispecies *Nitrosopumilus* (78%)
6	*Neptuniibacter caesariensis* (59%)	Marine Group I thaumarchaeote SCGC AAA799-E16 (85%)	Marine Group I thaumarchaeote SCGC AAA799-E16 (85%)	*Neptuniibacter caesariensis* (59%)
7	*Marinomonas* sp. D104 (59%)	*Marinomonas* sp. MWYL1 (59%)	*Marinomonas* sp. MWYL1 (59%)	*Marinomonas* sp. MWYL1 (58%)
8	Marine Group I thaumarchaeote SCGC AAA799-E16 (84%)	gamma proteobacterium BDW918 (58%)	gamma proteobacterium BDW918 (58%)	*Methylomicrobium kenyense* (59%)
9	*Enterovibrio norvegicus* (60%)	*Marinomonas ushuaiensis* (59%)	*Marinomonas ushuaiensis* (59%)	*Amphritea japonica* (58%)
10	*Enterovibrio calviensis* (59%)	*Marinomonas* sp. D104 (58%)	*Marinomonas* sp. D104 (59%)	*Endozoicomonas elysicola* (58%)

EctC	"*Ca.* Nitrosopumilus maritimus" SCM1	"*Ca.* Nitrosopumilus" sp. SJ	"*Ca.* Nitrosopumilus koreensis" AR1	"*Ca.* Nitrosopumilus" sp. AR2
1	*Nitrosopumilus maritimus* (100%)	Multispecies *Nitrosopumilus* (100%)	Multispecies *Nitrosopumilus* (100%)	Multispecies *Nitrosopumilus* (100%)
2	Multispecies *Nitrosopumilus* (82%)	Marine Group I thaumarchaeote SCGC AAA799-P11 (86%)	Marine Group I thaumarchaeote SCGC AAA799-P11 (86%)	Multispecies *Nitrosopumilus* (72%)
3	Marine Group I thaumarchaeote SCGC AAA799-P11 (83%)	*Nitrosopumilus maritimus* (82%)	*Nitrosopumilus maritimus* (82%)	Marine Group I thaumarchaeote SCGC AAA799-P11 (72%)
4	Marine Group I thaumarchaeote SCGC AAA799-N04 (83%)	Marine Group I thaumarchaeote SCGC AAA799-N04 (85%)	Marine Group I thaumarchaeote SCGC AAA799-N04 (85%)	*Nitrosopumilus maritimus* (71%)
5	Multispecies *Nitrosopumilus* (71%)	Multispecies *Nitrosopumilus* (72%)	Multispecies *Nitrosopumilus* (72%)	Marine Group I thaumarchaeote SCGC AAA799-N04 (71%)
6	*Sneathiella glossodoripedis* (60%)	*Bacillus fordii* (66%)	*Bacillus fordii* (66%)	*Hahella chejuensis* (63%)
7	*Desulfarculus baarsii* (59%)	*Hahella chejuensis* (60%)	*Hahella chejuensis* (60%)	*Hahella ganghwensis* (63%)
8	*Hahella chejuensis* (60%)	*Sneathiella glossodoripedis* (60%)	*Sneathiella glossodoripedis* (60%)	gamma proteobacterium BDW918 (61%)
9	*Nitrincola* sp. AK23 (60%)	*Paenibacillus pinihumi* (60%)	*Paenibacillus pinihumi* (60%)	*Desulfarculus baarsii* (57%)
10	*Vibrio* genomosp. F10 (61%)	*Nitrococcus mobilis* (62%)	*Nitrococcus mobilis* (62%)	*Sneathiella glossodoripedis* (57%)

EctD	"*Ca.* Nitrosopumilus maritimus" SCM1	"*Ca.* Nitrosopumilus" sp. SJ	"*Ca.* Nitrosopumilus koreensis" AR1	"*Ca.* Nitrosopumilus" sp. AR2
1	*Nitrosopumilus maritimus* (100%)	Multispecies *Nitrosopumilus* (100%)	Multispecies *Nitrosopumilus* (100%)	Multispecies *Nitrosopumilus* (100%)
2	Marine Group I thaumarchaeote SCGC AAA799-P11 (82%)	*Marine Group I thaumarchaeote SCGC AAA799-P11* (88%)	Marine Group I thaumarchaeote SCGC AAA799-P11 (88%)	Multispecies *Nitrosopumilus* (77%)
3	Multispecies *Nitrosopumilus* (81%)	*Marine Group I thaumarchaeote SCGC AAA799-N04* (87%)	Marine Group I thaumarchaeote SCGC AAA799-N04 (87%)	Marine Group I thaumarchaeote SCGC AAA799-P11 (66%)
4	Marine Group I thaumarchaeote SCGC AAA799-N04 (82%)	*Nitrosopumilus maritimus* (81%)	*Nitrosopumilus maritimus* (81%)	Marine Group I thaumarchaeote SCGC AAA799-N04 (77%)
5	Multispecies *Nitrosopumilus* (76%)	Multispecies *Nitrosopumilus* (77%)	Multispecies *Nitrosopumilus* (87%)	*Nitrosopumilus maritimus* (76%)
6	*Nitrosococcus halophilus* (59%)	*Nitrosococcus halophilus* (58%)	*Nitrosococcus halophilus* (58%)	*Nitrosococcus halophilus* (59%)
7	*Nitrosococcus oceani* (57%)	*Nitrosococcus watsonii* (56%)	*Nitrosococcus watsonii* (56%)	*Nitrosococcus watsonii* (57%)
8	*Paenibacillus senegalensis* (58%)	*Nitrosococcus oceani* (56%)	*Nitrosococcus oceani* (56%)	*Nitrosococcus oceani* (57%)
9	*Nitrosococcus watsonii* (56%)	*Oxalobacteraceae bacterium* IMCC9480 (58%)	*Oxalobacteraceae* bacterium IMCC9480 (58%)	*Paenibacillus senegalensis* (57%)
10	*Pusillimonas* sp. T7-7 (57%)	*Janthinobacterium* sp. Marseille (56%)	*Janthinobacterium* sp. Marseille (56%)	*Bacillus* sp. Ll(2012) (57%)

Amino acid sequence similarity of the ectoine biosynthetic enzymes from the four potential ectoine-producing "*Ca.* Nitrosopumilus" strains to other ectoine biosynthetic enzymes were assessed bioinformatically. The amino acid sequences of the EctA, EctB, EctC and EctD proteins of each strain were used as a query in BLAST searches. The top ten hits from each BLAST search are shown and the amino acid similarities are indicated in brackets.

Supplemental Table S2

Table S2. Distribution of mechanosensitive channels in potential ectoine-producing *Archaea*.

Organism (Accession no.)	MscS-type protein	MscL-type protein
"*Ca.* Nitrosopumilus maritimus" SCM1 (NC_010085.1)	ABX12867; ABX13231; ABX13669; ABX13233; ABX13238	not present
"*Ca.* Nitrosopumilus" sp. SJ (GCA_000328945.1)	WP_014963370; WP_014964111; WP_014963465	not present
"*Ca.* Nitrosopumilus koreensis" AR1 (GCA_000299365.1)	AFS80985; AFS81736; AFS81081	not present
"*Ca.* Nitrosopumilus" sp. AR2 (GCA_000299395.1)	WP_014965293; WP_014966054; WP_014964462	not present
Marine Group I thaumarchaeote SCGC AAA799-N04 (GCA_000722915.1)	WP_048069171; WP_048079437	not present
Marine Group I thaumarchaeote SCGC AAA799-P11 (GCA_000746685.1)	KFM18867; WP_048070662	not present
Marine Group I thaumarchaeote SCGC RSA3 (GCA_000746745.1)	WP_048069171; WP_048083714	not present
Marine Group I thaumarchaeote SCGC AAA799-E16 (GCA_000724145.1	WP_048088428	not present
***Methanothrix harundinacea* 6Ac**	WP_014587606; WP_014587910; WP_014586447; WP_014586460	AET64001; WP_048144789
***Methanothrix concilii* GP-6**	WP_013718893; AEB69207; WP_052297576; WP_013720440	WP_013718432
***Methanobacterium formicicum* DSM1535**	WP_048084718; WP_048072178; WP_004032010	not present
***Methanobacterium formicicum* BRM9**	WP_048084718; WP_048072178; WP_004032010	not present

The presence of potential mechanosensitive channels in ectoine-producing *Archaea* was analyzed bioinformatically. The amino acid sequences of MscS and MscL proteins from *Escherichia coli* were used as the search queries. Retrieved hits by the BLAST search of potential mechanosensitive channels are indicated by their NCBI accession numbers.

VII. Publication 7

The seventh manuscript is entitled 'Salt-sensitivity of σ^H and Spo0A prevents sporulation of *Bacillus subtilis* at high osmolarity avoiding death during cellular differentiation'. It was published in 'Molecular Microbiology' (doi: 10.1111/mmi.13304). Since this publication is the result of a side project of this dissertation and deals with a different topic, it is not part of the *Results* and *Discussion* sections, but can be found in the *Appendix* of this dissertation (Appendix II; page 168 *ff.*).

Widderich, N., Rodrigues, C.D.A., Commichau, F.M., Fischer, K.E., Ramirez-Guadiana, F.H., Rudner, D.Z. and Bremer, E. Salt-sensitivity of σ^H and Spo0A prevents sporulation of *Bacillus subtilis* at high osmolarity avoiding death during cellular differentiation. *Mol Microbiol*. (2016). 100: 108–124.

This paper addresses the question why sporulation is blocked in starving *Bacillus subtilis* cells that simultaneously face osmotic stress. It is demonstrated that spore formation is drastically impaired in starving high-salinity challenged *B. subtilis* cells and that the block occurs prior to asymmetric division. Further, it reveals that transcription dependent on σ^H and Spo0A is impaired and that association of σ^H with core-RNA-polymerase is reduced at high salinity. Additionally, it is demonstrated that cells induced to bypass the early block are impaired in later steps of the sporulation pathway and frequently undergo lysis prior to the formation of a mature spore indicating that the block in sporulation serves as a safeguard to avert osmotically stressed *B. subtilis* cells from committing to a developmental program they cannot complete and would die trying.

Personal contribution:
I conducted physiological experiments and performed immunoblot analysis and fluorescence microscopy experiments (together with Christopher Rodrigues). I further conducted pull-down experiments and enriched for suppressor strains. I analyzed the data (together with Christopher Rodrigues, David Rudner and Erhard Bremer) and prepared Figures and Tables (together with Christopher Rodrigues). I contributed to the writing of the manuscript together with Christopher Rodrigues, David Rudner and my PhD supervisor Erhard Bremer.

Discussion

In order to understand the importance of ectoine/5-hydroxyectoine synthesis in microbial osmostress adaptation, elucidation of the phylogenetic distribution of ectoine biosynthetic genes and their products was required prior to this dissertation. For the purpose of acquiring in-depth knowledge on the biochemistry of ectoine production, studies on the biochemical and kinetic properties of their key synthesizing enzymes, identification of amino acid side chains that are involved in catalysis in EctD and crucial for enzyme function and solution of the enzymes' crystal structure was of great importance. The aims of this dissertation dealt with these topics and its results will, with respect to afore accentuated cornerstones, be discussed in the following section.

I. Environmental role of ectoines

In their natural habitats, microorganisms frequently encounter a multitude of different stressful biotic and abiotic circumstances that can be detrimental for their growth and survival in a given ecosystem. One of the most important ecological challenges is alternating salinity/osmolarity due to its impact on almost all microorganisms (Csonka and Epstein, 1996; Kempf and Bremer, 1998; Lucht and Bremer, 1994; Pittelkow and Bremer, 2011). The permeability of the cytoplasmic membrane for water, as well as a considerable osmotic potential of the cellular cytoplasm, establish an intracellular hydrostatic pressure, the turgor, whose maintenance is considered to be critical for cell integrity (Booth and Blount, 2012; Booth *et al*., 2007; Bremer and Krämer, 2000; Koch, 1983; Wood, 2011). The desiccation of a natural habitat by different abiotic environmental factors, such as solar radiation and heat, leads to an increase in external salinity and osmolarity which inevitably triggers the efflux of water across the cytoplasmic membrane out of a bacterial cell (Csonka and Epstein, 1996; Kempf and Bremer, 1998; Lucht and Bremer, 1994; Pittelkow and Bremer, 2011). This water efflux impairs the maintenance of the cellular vital functions after the onset of plasmolysis that results from dehydration of the cytoplasm and from a collapse or a strong reduction of turgor (Booth and Blount, 2012; Booth *et al*., 2007; Bremer and Krämer, 2000; Koch, 1983; Wood, 2011). To avoid dehydration of the cytoplasm and a collapse of turgor, many microorganisms have evolved mechanisms that counteract the diffusion of water out of the bacterial cell under high osmolarity in order to balance the osmotic gradient (Galinski and Trüper, 1994; Pittelkow and Bremer, 2011).

In the initial phase of their osmotic adjustment process many bacteria accumulate large quantities of potassium (K^+) ions via specialized transport systems. In a second step, the cell replaces these K^+-ions by the uptake or *de novo* synthesis of a selected class

of low-molecular organic compounds, the so-called “compatible solutes”, in order to reduce the ionic strength of the cytoplasm (Csonka and Epstein, 1996; Kempf and Bremer, 1998; Lucht and Bremer, 1994; Pittelkow and Bremer, 2011). Two of the most widely used compatible solutes by members of the *Bacteria* are the tetrahydropyrimidines ectoine and its hydroxylated derivative 5-hydroxyectoine (Fig. 4). Both substances are amino acid derivatives and, as all compatible solutes, osmotically active, low molecular compounds that can be accumulated to very high intracellular concentrations without disturbing cellular metabolism and vital functions (Bursy *et al*., 2007; da Costa *et al*., 1998; Galinski, 1993; Grant, 2004; Inbar and Lapidot, 1988; Ono *et al*., 1999; Peters *et al*., 1990; Ventosa *et al*., 1998).

Although both, ectoine and 5-hydroxyectoine, serve as effective osmoprotectants, 5-hydroxyectoine possesses stabilizing properties for macromolecules that are often superior to those exhibited by ectoine (Borges *et al*., 2002; Kurz, 2008; Lippert and Galinski, 1992; Schnoor *et al*., 2004). Ectoine and 5-hydroxyectoine are commercially used as stabilizers for proteins but foremost in cosmetics and skin care products; medical applications of these two compounds are actively pursued as well [bitop AG, Witten; http://www.bitop.de/]. Hence, there is a considerable interest to further understand the properties of the enzymes involved in the ectoine and 5-hydroxyectoine biosynthetic route.

Therefore, a comprehensive study on the phylogeny, biochemistry and structure of the key enzymes in ectoine (EctC) and hydroxyectoine (EctD) biosynthesis was conducted in the course of this dissertation. I demonstrated that ectoine and hydroxyectoine genes are widely distributed in bacteria and that they are also present in a few archaea indicating that synthesis of these chemical chaperones is used by many microorganisms to fend off the detrimental effects caused by high osmolarity. Moreover, representative ectoine hydroxylases and two ectoine synthases have been biochemically characterized. Finally, a crystal structure of EctD has been solved in its apo-form, in complex with its co-factors and its reaction product in collaboration with Dr. Sander Smits (University of Düsseldorf). In addition, a detailed expression and purification protocol for the ectoine synthase from the cold-adapted marine bacterium *Sphingopyxis alaskensis* has been described in collaboration with Dr. Sander Smits (University of Düsseldorf) that identifies EctC to form dimers in solution and deals with crystallization trials and preliminary X-ray diffraction data of the EctC enzyme.

II. Phylogeny of ectoine and hydroxyectoine biosynthesis

The ectoine biosynthetic enzymes EctB (EC 2.6.1.76) and EctA (EC 2.3.1.178) have enzymatic counterparts in a variety of metabolic pathways (Ono *et al*., 1999; Peters *et al*., 1990), but the ectoine synthase EctC (EC 4.2.1.108) is unique, meaning that analogs of this enzyme are not apparent in other biochemical routes. EctC and EctD are considered as the key enzymes in ectoine and hydroxyectoine biosynthesis, respectively, and can therefore be used as the diagnostic enzymes in screens for production of these important chemical chaperones. Data reported in this dissertation illustrate that the ability to produce ectoines is widely distributed among prokaryotes (Fig. 9) (Publication 3 and 6) (Widderich *et al*., 2016a; Widderich *et al*., 2014a).

At the search date (08.07.2015), 27.232 genome sequences of members of the *Bacteria* and 557 genome sequences of *Archaea* were represented in the JGI microbial database. Among these microbial genome sequences, in total 27.789, 1297 hits to EctC-type proteins were retrieved. After removal of redundant entries (e.g. there were 181 genomes of strain of *Vibrio cholerae* represented of which each possessed an *ect* gene cluster), the final data set comprised 723 EctC-type proteins.

The domain of the *Bacteria* represents the far most dominant group of the potential ectoine producers, counting about 711 representatives distributed in about 19 bacterial phyla. In *Archaea*, only a few representatives (12) were identified that possess ectoine biosynthetic genes. These belong to three different archaeal families. Interestingly, ectoine biosynthesis genes have not yet been identified in *Eukarya*. Based on the heterogeneous distribution of ectoine biosynthetic genes, one can speculate that the ectoine biosynthesis gene cluster has been initially developed and evolved in *Bacteria*, has then been established due to its important stress-protectant functions and has finally been distributed among microorganisms via lateral gene transfer events. In line with this conclusion, the distribution of the ability to synthesize ectoine, as indicated by the presence of an *ectC* gene in a given genome sequence, appeared to extend mostly to members of the *Proteobacteria*, *Firmicutes*, and *Actinobacteria* (Publication 3 and 6) (Widderich *et al*., 2016a; Widderich *et al*., 2014a). In addition, the EctC sequences cluster mainly within the taxonomic subgroups of these bacterial phyla and their branching order occurs largely in parallel with the different corresponding taxonomic units of their microbial hosts down to the order level. This suggests a long co-evolution of the ectoine biosynthetic genes in the various bacteria.

This hypothesis is further strengthened by having a closer look onto the natural habitats and ecological niches of all potential ectoine-producing microorganisms. In this context, it appears that those microorganisms that possess ectoine biosynthesis genes mainly colonize marine and terrestrial habitats. Hence, there is an ideal basis for lateral

gene transfer events of the ectoine gene cluster. As an example, it can be speculated that the marine ammonia-oxidizing archaeal isolate *Nitrosopumilus maritimus* (Konneke *et al*., 2005) received its *ect*-cluster via horizontal gene transfer. This is probably also the case for the other archaeal ectoine producers, since the 12 EctC proteins derived from Archaea are found in three different locations in the ectoine synthase-based phylogenetic tree (Fig. 9) and thus presumably received their ectoine biosynthetic genes by separate transfer events. Another example for potential horizontal gene transfer events of the *ect*-cluster can be found in the δ-*proteobacteria* phylum, since all identified representatives exhibit *ectC* genes highly dispersed on the phylogenetic tree of ectoine synthases (Fig. 9). This means that their *ectC* gene products significantly differ in their amino acid sequence and were therefore likely acquired via single and independent lateral gene transfer events from different donors. Beside the mentioned examples, there are also other bacterial phyla, such as α- and γ-*proteobacteria*, which are apparently distributed on the phylogenetic tree and thus reinforce the above-made hypothesis. In remarkable contrast, the phylum of *Actinobacteria*, which represents the largest group of potential ectoine producers, is largely cohesive on the phylogenetic tree, suggesting that the capability to synthesize ectoines evolved early within this phylum. However, one should handle these data and the resulting conclusions and speculations with care, since the used databases are biased comprising only a small fraction (mainly cultivatable) of microorganisms and thus might not reflect the genuine situation in nature.

A large part of the potential ectoine producers (~ 60%) also possesses an *ectD* gene within its genome and is thus presumably able to synthesize hydroxyectoine, as well. *Bona-fide* EctD-type proteins can be identified by the presence of a strictly conserved signature sequence (Bursy *et al*., 2007; Reuter *et al*., 2010) that allows distinguishing ectoine hydroxylases from related proteins. In about 70 % of the potential hydroxyectoine producers, the *ectD* gene can be found as part of the *ect*-cluster, downstream of *ectC*. Again, the phylum of the *Actinobacteria* represents the largest group among all potential hydroxyectoine producers. Looking at the distribution of *ectD* in its entirety, one recognizes that the ectoine hydroxylase is in particular present in terrestrial microorganisms, whereas marine microbes build a minority of the potential hydroxyectoine producers. This observation can be easily explained when one considers the fact that soil-dwelling microorganisms, other than such from marine habitats, also often encounter, besides osmotic stress, extremes in temperature due to solar radiation and heat. Thereby, temperatures can rapidly rise locally up to 50° C, an increase aquatic microorganisms quite rarely experience. In fact, ectoine and hydroxyectoine do not significantly differ in their osmoprotectant properties, but

distinctions with respect to their thermo-protective characteristics have been reported (Ablinger *et al*., 2012; Bursy *et al*., 2008; Kuhlmann *et al*., 2008a; Kuhlmann *et al*., 2011; Garcia-Estepa *et al*., 2006; Tanne *et al*., 2014). With this in mind, one can speculate that the capability to convert ectoine into hydroxyectoine is therefore predominantly present in terrestrial microorganisms, since hydroxyectoine is superior over ectoine under thermostress conditions and might thus provide a selective advantage under these specific circumstances. In line with this hypothesis is the fact that the ectoine hydroxylase is not always part of the *ect* gene cluster but encoded at another position in the genome. Here, its expression might be differentially regulated compared to the *ect* operon allowing transcription of *ectD* only under environmental circumstances its gene product is beneficial. Moreover, it is of no surprise that strict anaerobes predicted to produce ectoine, such as members of the *Methanosaeta* and *Methanobacterium* genera, do not possess *ectD*, since the ectoine hydroxylase (EctD) is an oxygen-dependent enzyme (Bursy *et al*., 2007). Hence, these organisms would not be able to catalyze the hydroxylation of ectoine to yield 5-hydroxyectoine.

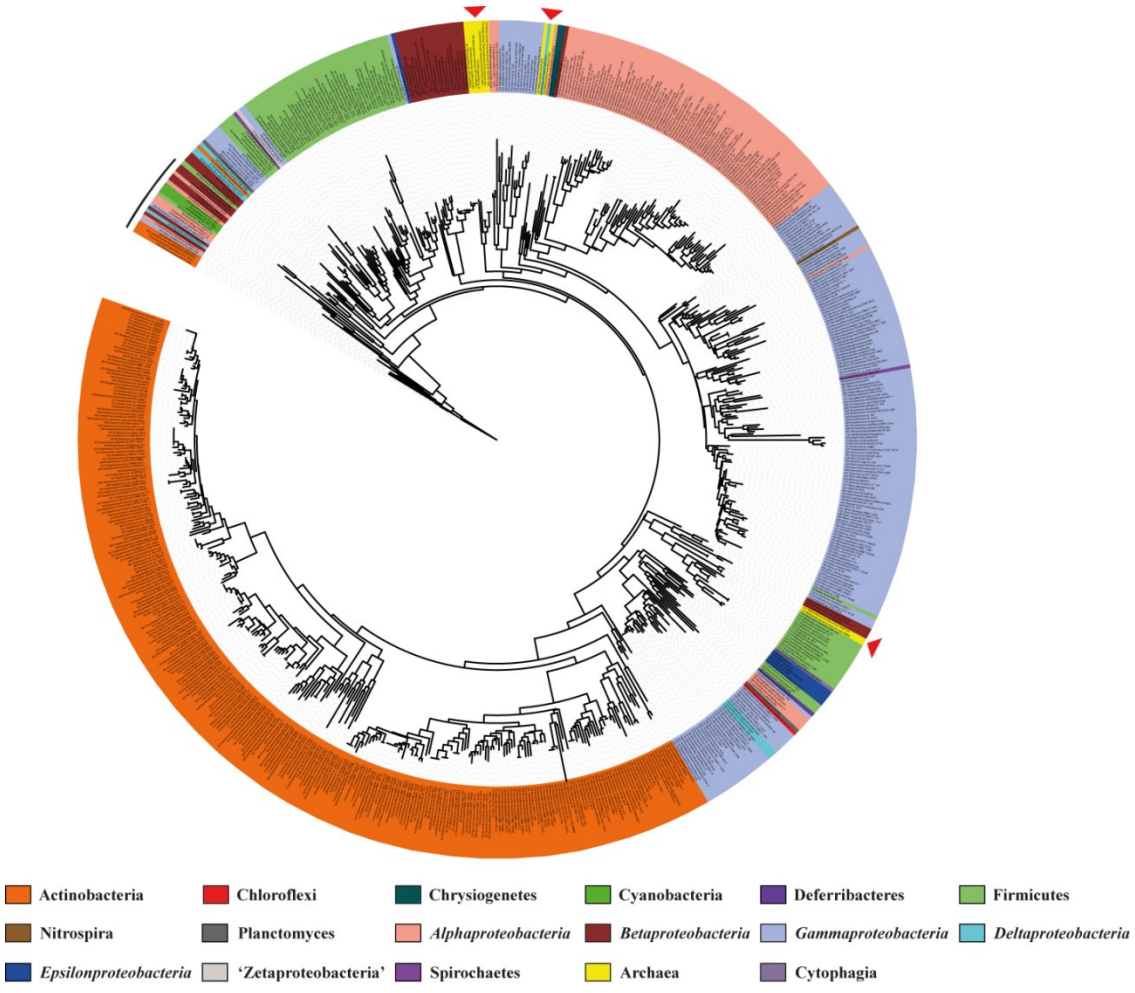


Fig 9. Phylogenetic tree of EctC-type proteins. Based on an amino acid sequence alignment of 723 EctC-type proteins, a rooted phylogenetic tree is shown constructed with the iTOL software (Letunic and

Bork, 2011). The different phyla are represented by different colors. The three regions in the phylogenetic tree populated by archaeal EctC proteins are highlighted by red arrowheads. The position of those 24 orphan EctC-type proteins originating from bacteria that lack identifiable *ectAB* genes are marked above the phylogenetic tree by a black bar. Modified from Widderich *et al.* (Widderich *et al*., 2016a).

Beside the typical ectoine biosynthetic genes, several additional genes have been identified in the course of this dissertation that are frequently associated with the *ect*-cluster or locate in its close vicinity. The most abundant gene, called *ask_ect*, encodes a specialized aspartokinase (Stöveken *et al*., 2011). This gene product catalyzes the reaction from L-aspartate to aspartate-β-phosphate, which is an important intermediate in diverse metabolic pathways and a precursor molecule for the production of ectoine. Since phosphorylation of metabolic intermediates is quite energy-consuming, kinases are often strictly feedback-regulated by transcriptional, posttranscriptional and posttranslational control mechanisms to ensure that there is no over- or undersupply of metabolic substrates; a fact that is also true for aspartokinases. These enzymes are ubiquitously found in microorganisms and several aspartokinases with distinct regulatory features have been shown to be often present in the same bacterial cell (Lo *et al*., 2009). This is also true for the ectoine/hydroxyectoine producing bacterium *Pseudomonas stutzeri* A1501. Recently, comparative studies on the biochemistry of the specialized aspartokinase Ask_Ect and the anabolic aspartokinase LysC revealed distinct feedback inhibition profiles by metabolites of the two proteins. The cellular pool of L-aspartate-β-semialdehyde is a potential bottleneck for the massive ectoine synthesis initiated under high osmolarity (Bestvater *et al*., 2008; Stöveken *et al*., 2011). Acquisition of a specialized aspartokinase might avoid such a metabolic constraint and therefore explains its presence in a substantial sub-group of the ectoine producing microbes (Widderich *et al*., 2014a; Widderich *et al*., 2016a).

Another gene often identified to locate in the direct neighborhood of the ectoine biosynthetic gene cluster is called *ectR*. Its corresponding gene product belongs to the family of MarR-type transcriptional regulators. In previous studies, it has already been characterized as a negative regulator of the *ect*-operon coordinating transcription of the whole ectoine gene cluster (Mustakhimov *et al*., 2012; Mustakhimov *et al*., 2010). However, elevated transcription of the ectoine biosynthetic genes in an *ectR* mutant remained osmotically inducible (Mustakhimov *et al*., 2010). In *Methylomicrobium alcaliphilum*, the EctR operator overlaps the -10 sequence of the *ect* promoter and EctR might also regulate the expression of its own structural gene; but the latter regulatory feature does not always seem to exist (Mustakhimov *et al*., 2012). Moreover, it is currently still unknown which environmental or cellular cues dictate the binding to or the release of EctR from its operator sequence.

Nevertheless both, EctR and Ask_ect, perform further important functions/roles in the biosynthesis of ectoines by on the one hand strictly regulating transcription of the *ect*-operon and on the other hand catalyzing an additional step in the ectoine biosynthetic route and thus optimize cellular production of ectoines. The fact that both are only present in a minority of the potential ectoine producing microorganisms, and here predominantly in *Proteobacteria*, suggests that these genes have been developed or established posterior to ectoine biosynthesis.

In some potential ectoine producers several further genes have been identified in the course of this dissertation to locate in direct vicinity to the ectoine biosynthetic genes that encode potential transport systems. These are often ABC transport systems whose binding proteins frequently possess high sequence identity to already characterized ectoine binding proteins, such as EhuB from *Sinorhizobium meliloti* (Hanekop *et al*., 2007). Hence, these genes probably mediate the uptake of ectoines under osmotic stress or might act as recapturing advice for ectoine that is synthesized *de novo* and subsequently released by the stressed cells as observed for proline in osmotically stressed *B. subtilis* cells (Hoffmann *et al*., 2012). For the ectoine transport system TeaABC, it has also already been assumed that its main physiological role is to primarily recover ectoine leaking through the cytoplasmic membrane (Grammann *et al*., 2002). Moreover, genes encoding potential MFS- and SSS-transporters have also been identified that show a high degree of sequence identity to the OpuE transport system of *B. subtilis* known to be responsible for the uptake of proline (Moses *et al*., 2012; von Blohn *et al*., 1997). Considering the high structural relatedness of ectoine and proline, one can yet assume that these transporters also mediate the import of ectoines since only minor mutational changes in the binding motif of these transporters might be sufficient to bind ectoine instead of proline. Aside from these typical transport systems, genes encoding Msc-type proteins have also been identified in the neighborhood of ectoine biosynthetic genes that encode potential mechanosensitive channels. These mechanosensitive channels act in both, *Bacteria* and *Archaea*, as safety valves to withstand rapid osmotic down-shifts and are widely considered as non-specific pores that open under hypo-osmotic conditions to quickly release osmotically active solvents from the cytoplasm thereby preventing cellular rupture and lysis (Booth, 2014; Booth and Blount, 2012; Kloda and Martinac, 2002; Wilson *et al*., 2013). One prominent example is the ectoine biosynthetic gene cluster of the thaumarchaeote *N. maritimus*. As demonstrated within this dissertation in publication 6 (Widderich *et al*., 2016a) the *mscS* gene is part of the *ect*-operon and its expression is induced upon osmotic stress indicating a kind of pre-adaptation to a likely occurring down-shift in environmental osmolarity. As its functionality was also experimentally confirmed in

publication 6, the *ect*-operon of *N. maritimus* appears to be a rather sophisticated genetic device that counteracts the detrimental effects of osmotic fluctuations on cell physiology, which likely occur in its natural ocean habitat. The described *ect*-operon thus allows the cell to deal with increased salinity through the production of the osmostress protectants ectoine and hydroxyectoine and at the same time already prepares the osmotically stressed cell for a potential occurring osmotic down-shift through the preventive synthesis of mechanosensitive channels.

An interesting finding while the phylogenetic diagnosis of ectoine biosynthesis was the identification of a restricted number of microorganisms that possess *ectC* type genes but lack the corresponding *ectAB* genes in their genomes (Widderich *et al*., 2016a; Widderich *et al*., 2014a). Closer inspection of these orphan *ectC* genes revealed that their taxonomic affiliation is rather diverse and that they all cluster together forming a distinct branch at the root of the EctC-protein tree (Publication 6) (Widderich *et al*., 2016a). One of these orphan *ectC* gene products has already been studied in *Pseudomonas syringae* pv. syringae strain B728a prior to this dissertation and the authors found that ectoine was only produced under osmotic stress conditions when surface-sterilized leaves of the host plant *Syringa vulgaris* were added to the bacterial culture (Kurz *et al*., 2007). In addition, functional complementation experiments revealed that the corresponding ectoine synthase was only modestly active and supply of *N*-γ-acetyl-L-2,4-diaminobutyrate, the natural substrate of *bona fide* ectoine synthases, did not result in ectoine synthesis, although it has been shown that it is imported by the cells (Kurz *et al*., 2007). Hence, the functional relevance of these orphan *ectC* genes is still illusive. Microorganisms possessing them might be actual ectoine producers that have to rely on an environmental supply of ectoine precursor molecules as suggested by Kurz *et al.* (Kurz *et al*., 2007); a hypothesis that seems reasonable when the precursor molecule is available, as it might be the case for microorganisms that occur in close vicinity to plants or live in symbiosis, commensalism or antibiosis. Alternatively, these orphan *ectC* genes might be remnants of evolution of a previously fully intact ectoine biosynthetic pathway or might have evolved towards biochemical activities other than ectoine synthesis. Another possibility would be that these orphan *ectC* genes are evolutional ancestors of ectoine biosynthesis. A hypothesis that is strengthened by the fact that all solitary *ectC* genes form a distinct branch at the root of the tree as mentioned above.

The elucidation of the phylogenetic distribution of ectoine and hydroxyectoine biosynthesis is not only the first detailed analysis of the dissemination of compatible solutes contributing to the understanding of the importance of chemical chaperones such as ectoines in microbial stress resistance, but also puts forward an in-depth

biochemical study of their synthesizing enzymes on a broad basis. Based on these results, the determination of the biochemistry and kinetics of the key enzymes in ectoine and hydroxyectoine biosynthesis is discussed in the following section.

III. Biochemistry and kinetics of the ectoine synthase and the ectoine hydroxylase

To further understand the molecular details and physiological role of the key enzymes in ectoine and hydroxyectoine biosynthesis, it is of importance to elucidate the biochemical and kinetic properties of the ectoine synthase and the ectoine hydroxylase. Therefore, special emphasis was placed within this dissertation on the biochemical and kinetic characterization of both the ectoine synthase and the ectoine hydroxylase. The phylogenetic distribution discussed in the previous section provided the basis to sample a sequence space of EctD- and EctC-type proteins.

In terms of the biochemistry of EctD, the salient biochemical and kinetic properties of in total seven ectoine hydroxylases have, partially based on previous studies (Bursy *et al*., 2007; Pittelkow, 2011; Widderich, 2012), been characterized within this dissertation (Widderich *et al*., 2016a; Widderich *et al*., 2014a; Widderich *et al*., 2014b). With respect to EctD, enzymes were chosen that were widely distributed on the phylogenetic tree of ectoine biosynthesis (Fig. 9) and additionally were derived from microorganisms that colonize habitats with extremes in pH (*Alkalilimnicola ehrlichii, Acidiphilium cryptum*), in temperature (*Sphingopyxis alaskensis, Paenibacillus lautus* Y412MC10) and in salt content (*Halomonas elongata*). Two further EctD-type proteins from non-extremophiles were also investigated. The first enzyme is derived from the γ-*proteobacterium Pseudomonas stutzeri* and the second from the crenarchaeon *Nitrosopumilus maritimus*. The *N. maritimus* EctD protein was chosen since it was to this date the only Archaeon known so far harboring an *ectD* gene. The *P. stutzeri* EctD enzyme was analyzed due to our interest in its additional *ask_ect* gene within the ectoine biosynthesis cluster (Stöveken *et al*., 2011). Their biochemical analyses have been published in two original manuscripts (Publications 3 and 6) (Widderich *et al*., 2016a; Widderich *et al*., 2014a). With respect to EctC, enzymes were chosen the same way as for EctD. Here, initial biochemical and/or kinetic properties of two ectoine synthases (Publications 5 and 6) (Widderich *et al*., 2016a; Kobus *et al.*, 2015) have been characterized within this dissertation. These also derived from *S. alaskensis* (publication 5), the cold-adapted member of the *Bacteria* as outlined above, as well as from the thaumarchaeote *N. maritimus* (publication 6), which has also been described

above. Their analyses have been reported in two original manuscripts (Publications 5 and 6) (Widderich *et al*., 2016a; Kobus *et al.*, 2015).

The seven ectoine hydroxylases analyzed in the course of this dissertation and the two characterized in previous studies (Bursy *et al*., 2008; Bursy *et al*., 2007) were highly similar with respect to their kinetic parameters and showed comparable catalytic efficiencies (Publications 3 and 6). The same was true for the kinetic parameters of the tested EctC enzyme (Publications 6) and those characterized and published previously (Moritz *et al*., 2015; Ono *et al*., 1999; Reshetnikov *et al*., 2006). The fact that both, EctD and EctC, possessed K_m values for their substrates in the low millimolar range can be easily explained by the fact that the intracellular pool of ectoines in response to osmotic stress may probably reach up to 500 mM, as it is the case for proline in *B. subtilis* cells facing severe osmotic stress conditions (Brill *et al*., 2011; Hoffmann *et al*., 2013). Hence, the intracellular availability and provision of precursor molecules, e.g. *N*-γ-ADABA, needs to be high (in the millimolar range) for this purpose as well. Thus, there is no need for high affinity binding of the ectoine biosynthesis key enzymes for their substrates, as it is the case for compatible solute import systems (often in the µM range) that need to trap their ligands from the environment occurring in yet residual concentrations (Hanekop *et al*., 2007; Lecher *et al*., 2009; Kuhlmann *et al*., 2008b).

With respect to their salient biochemical properties, the ectoine hydroxylases from different organisms also appeared to have quite similar characteristics since they possessed similar temperature and pH optima as well as comparable stimulation by different salts (Publications 3 and 6) (Bursy *et al*., 2008; Bursy *et al*., 2007; Widderich *et al*., 2016a; Widderich *et al*., 2014a). However, an interesting observation was made with respect to the temperature stability of the investigated EctD-enzymes. Again, most enzymes showed very similar activities after incubation at increasing temperatures, but two EctD proteins stood out from the others (Publication 3) (Widderich *et al*., 2014a). Notably, these were the ectoine hydroxylases from both thermo-extremophilic microorganisms. This observation could have been expected for the EctD enzyme from *Paenibacillus lautus* Y412MC10 which is a thermotolerant microorganism isolated from a hot mud-whole in the Yellowstone National Park and by considering that the intracellular accumulation of hydroxyectoine increases the thermo-tolerance of microorganisms (Bursy *et al*., 2008; Garcia-Estepa *et al*., 2006). Since *P. lautus* frequently encounters high temperatures in its natural habitat, a temperature-stable version of EctD might provide selective growth advantage due to optimized synthesis of hydroxyectoine. In contrast, the fact that the ectoine hydroxylase from the psychrophilic bacterium *Sphingopyxis alaskensis* isolated from the Gulf of Alaska showed improved temperature stability as well was quite surprising and thus, speculations on the EctD

enzyme of *S. alaskensis* are tougher. One explanation of increased thermo-tolerance of *Sa*EctD might be lateral gene transfer of *ectD* from a more thermophilic organism, but this remains unclear and speculative so far.

In contrast to the studied ectoine hydroxylases, inspection and comparison of the EctC proteins characterized in the course of this dissertation (Publications 5 and 6) (Widderich *et al*., 2016a; Kobus *et al*., 2015) and the three analyzed in previous studies (Moritz *et al*., 2015; Ono *et al*., 1999; Reshetnikov *et al*., 2006) revealed major differences with respect to their salient biochemical properties. It seems that these enzymes reflect well the natural habitats of their hosts indicating that the ectoine synthases have adapted to the lifestyle of these microorganisms. This hypothesis is in line with asumptions made prior and posterior to this dissertation (Moritz *et al.,* 2015; Widderich *et al.,* 2016d) and is strengthened by the fact that EctC proteins originating from halophilic organisms showed a stronger stimulation of enzyme activity and tolerance to different salts than those derived from non-halophilic microorganisms (Publications 6) (Ono *et al*., 1999; Moritz *et al*., 2015). A finding not observed for the studied ectoine hydroxylases, which all exhibited very similar responses to the addition of salts.

Significant differences with respect to optimal pH and tolerated pH ranges were neither identified for EctD nor for EctC proteins. However, taking into account that the intracellular pH value is kept quite constant and similar by microbial pH homeostasis regardless of the surrounding pH (Krulwich *et al*., 2011), this observation is readily understandable.

Collectively, seven ectoine hydroxylases and two ectoine synthases have been characterized within the course of this dissertation. The presented data suggest, in connection with previous observations (Ono *et al.,* 1999; Moritz *et al.,* 2015), that the ectoine synthase has co-evolved with its host microorganisms and is thus well adapted to their lifestyles, whereas the ectoine hydroxylase is not, despite the fact that the studied EctD enzymes derived from various microorganisms thriving highly different habitats. A hypothesis further underlined by the observation that ectoine hydroxylases possess, in comparison to ectoine synthases, a significant higher degree of sequence identity indicating that EctD-type proteins are more closely related to each other than EctC-type enzymes. Finally, one can assume from the results of this dissertation that EctC, the key enzyme for ectoine biosynthesis, has been developed early within the tree of life, has co-evolved with its host organisms and has further been distributed by horizontal gene transfer events. In contrast, one can surmise that EctD, the key enzyme of hydroxyectoine biosynthesis, has likely developed to a later stage and might yet be at the beginning of an evolutional process.

The data discussed in this section contribute significantly to the understanding of the synthesis of ectoines and sets, together with the bioinformatic analysis, a solid basis for new crystallographic approaches by using enzymes that might be better suited for these trials due to e.g. increased protein stability. Crystallization trials for EctC and solution of the crystal structure of EctD as well as the identification of amino acid residues important for ligand binding and catalysis within the catalytic core of EctD is discussed in the following section.

IV. Crystal structure and site-directed mutagenesis of the ectoine hydroxylase and crystallization trials of the ectoine synthase

To fully understand the biochemistry of the key enzymes in ectoine and hydroxyectoine biosynthesis, as well as to unravel their mediated reaction mechanisms, the solution of their crystal structures is indispensable.

The ectoine hydroxylase EctD has previously been shown to be member of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily (Bursy *et al*., 2007; Reuter *et al*., 2010), a subfamily of the large cupin superfamily. Moreover, a crystal structure of EctD from the moderate halophile *V. salexigens* has also already been reported in literature (Reuter *et al*., 2010). This structure revealed the positioning of the iron co-factor within the active site of the ectoine hydroxylase, but it contained neither the substrate ectoine nor the co-substrate 2-oxoglutarate. Hence, prior to this dissertation, the coordination of the ligands within the active site, as well as the catalyzed enzyme reaction mechanism, remained still unclear.

In terms of EctC, considerations based on bioinformatic analysis of microbial genome databases (Dunwell *et al*., 2001; Agarwal *et al*., 2009; Uberto and Moomaw, 2013) suggest that the ectoine synthase belongs to the cupin superfamily, as well. This large group of proteins possesses a conserved ß-barrel structural scaffold and performs versatile enzymatic and non-enzymatic functions (Dunwell *et al*., 2001). Proteins of the cupin family often harbor transition metals such as iron, copper, zinc, manganese, cobalt, or nickel, which are of importance for catalysis and enzymatic function (Agarwal *et al*., 2009; Dunwell *et al*., 2001; Uberto and Moomaw, 2013). Trials to solve the structure of EctC have also been reported (Witt, 2012), but a high-resolution crystal structure has not been published to date. Hence, coordination of the ligands within the active site of EctC and its catalyzed reaction mechanism remain illusive. Moreover, it was not known whether the ectoine synthase is also a metal-dependent enzyme as it is the case for many, but not all, enzymes of the cupin family, since the few EctC proteins

studied so far were not analyzed in this respect (Moritz *et al*., 2015; Ono *et al*., 1999; Reshetnikov *et al*., 2006).

Therefore, crystallographic trials were started within this dissertation in collaboration with Dr. Sander Smits (University of Düsseldorf) for both, the EctC and the EctD proteins. These trials were based on the use of enzymes that have been identified to show higher enzyme stability during their biochemical characterization as outlined in the previous sections.

These new approaches resulted with respect to EctD in four further solved crystal structures of the ectoine hydroxylase: i) one from the moderate halophile *V. salexigens* in its apo-form, ii) and three from the cold-adapted ultrabacterium *S. alaskensis* in its apo-form, in complex with the iron catalyst and in a form containing the iron catalyst, the co-substrate 2-oxoglutarate and the product of the EctD-catalyzed enzyme reaction, 5-hydroxyectoine. These structures have been deposited in the protein data bank (http://www.rcsb.org/pdb/home/home.do) and were published in three manuscripts (Publications 2 – 4) (Hoeppner *et al*., 2014b; Widderich *et al*., 2014a; Hoeppner *et al*., 2014a).

In terms of EctC, the crystallization trials resulted in well diffracting crystals of the *S. alaskensis* EctC protein (Publication 5) (Kobus *et al*., 2015) that provide a good basis to solve the crystal structure of the ectoine synthase with all its ligands.

The ectoine hydroxylase is a member of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily. Members of this superfamily share similar overall structural folds and enzyme mechanisms, but they can catalyze a diverse set of oxidation reactions (Aik *et al*., 2012; Hangasky *et al*., 2013; Hausinger, 2004; Kundu, 2012; Wong *et al*., 2013). The crystal structures of EctD from the extremophilic bacteria *V. salexigens* (Publication 3) (Reuter *et al*., 2010) and *S. alaskensis* (Publication 4) also share a high degree of identity in their overall fold and are most closely related to the human hydroxylases PhyH and PHYD1A (McDonough *et al*., 2005; Zhang *et al*., 2011) and to the microbial halogenases SyrB2, CytC3, and CurA (Blasiak *et al*., 2006; Khare *et al*., 2010; Wong *et al*., 2009). Moreover, the crystal structures adhere to the common protein fold of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily, in which several antiparallel β-sheets form one side of a barrel-like structure (Aik *et al*., 2012; Hangasky *et al*., 2013). A variety of quaternary conformations are reported for enzymes belonging to this superfamily (Aik *et al*., 2012; Hangasky *et al*., 2013; Hausinger, 2004; Knauer *et al*., 2012) and EctD has been previously reported to be a monomer (Bursy *et al*., 2007; Reuter *et al*., 2010). However, within this dissertation it has been demonstrated that all studied ectoine hydroxylases form homodimers in solution being oriented in a head-to-

tail fashion as observed by crystal analysis (Fig. 10 A) (Publications 3 and 4). The crystallographic analysis of the *Sa*EctD protein in complex with its ligands reported in publication 4 (Hoeppner *et al*., 2014b) revealed that three residues forming a conserved 2-His-1-Asp facial triad bind iron (Aik *et al*., 2012; Hangasky *et al*., 2013; Hausinger, 2004; Kundu, 2012; Straganz and Nidetzky, 2006), as it was also observed before for the *Vs*EctD protein (Reuter *et al*., 2010). In this connection, it is worth mentioning that both, the solved apo-EctD structures and the structures with bound iron catalysts, completely matched in their overall fold (Publications 3 and 4) (Reuter *et al*., 2010) indicating that the ectoine hydroxylase adopts its tertiary conformation prior to the binding of the iron co-factor.

Through the solution of the *Sa*EctD crystal structure in complex with all ligands and by molecular dynamics simulations of the *Vs*EctD protein several amino acids have been identified in the course of this dissertation that contribute to ligand binding. Key contributions to the precise positioning of these ligands are mediated by residues present on an extended α-helix and a flanking short β-strand, which contain most of the residues making up the strictly conserved signature sequence of EctD-type hydroxylases (Bursy *et al*., 2007; Reuter *et al*., 2010). Alignments of all ectoine hydroxylase amino acid sequences available in the database revealed that each of these side chains is highly conserved among EctD-type proteins, reinforcing their crucial role in catalysis. In connection with comprehensive site-directed mutagenesis experiments performed in the course of this dissertation (Publications 1 and 4), their functional role was additionally probed and demonstrated their importance, thereby also explaining their high degree of conservation.

Overall, the combination of molecular dynamics simulations, solution of multiple EctD crystal structures and a comprehensive set of site-directed mutagenesis experiments within the course of this dissertation have demonstrated that an intricate network of interactions is present in the catalytic core of EctD occurring between the iron catalyst, the co-substrate 2-oxoglutarate, the substrate ectoine and their binding amino acid side chains (Fig. 10 B). This complex network of interactions thereby mediates the positioning of the substrates within the active site in such a fine-tuned way that the hydroxylation of ectoine occurs in a highly region-selective and stereo-specific manner at the C5-carbon atom within the pyrimidine ring of ectoine.

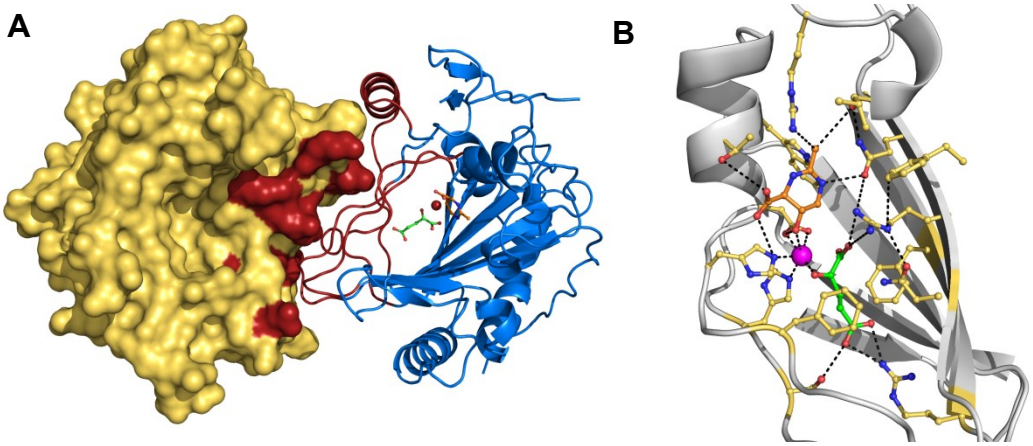


Fig 10. Crystal structure of the ectoine hydroxylase EctD. (A) Representation of the monomer/monomer interface of the (*Sa*)EctD dimer in complex with iron(II), 2-oxoglutarate and 5-hydroxyectoine. One monomer is shown in surface representation, whereas the other monomer is represented in ribbon format. Residues and regions that interact in the dimer are highlighted in red. The ligands are depicted in the monomer shown in ribbon format. The iron catalyst is represented as a red sphere; the co-substrate 2-oxoglutarate (green) and 5-hydroxyectoine are shown in ball and stick representation (orange). (B) Active site of the EctD crystal structure highlighting the intricate network of interactions that mediates the binding of the iron catalyst (magenta sphere), the 2-oxoglutarate co-substrate (green sticks) and the EctD-catalyzed enzyme reaction product, 5-hydroxyectoine (orange sticks). Amino acids involved in ligand binding are represented as yellow sticks. Modified from Hoeppner *et al.* (Hoeppner *et al*., 2014b).

A comparison of the different EctD crystal forms among themselves indicates that the ectoine hydroxylase possesses a highly-flexible region close to its carboxy-terminus (Publications 4) suggesting that the EctD crystal structures were solved in different catalytic states. Similar movements that correspond to a lid-like structure have been already identified in other members of the cupin family (Giraud *et al*., 2000; Xu *et al*., 2014). It can be speculated that this flexible loop also adopts such a lid-like structure posterior to the binding of the ligands, which closes the active site from its surroundings. Lid-closure might then accompany a compaction of the proteins' active site and an out-crowding of water molecules from the catalytic core. However, since loop-movements have not been detected for all the solved structures, this hypothesis remains speculative.

Being successful with the crystallization trials for EctD, a similar approach was conducted for the solution of the crystal structure of the ectoine synthase by using enzymes with increased protein stability. This approach seems to turn out to be effective as well. Crystallization trials performed in collaboration with Dr. Sander Smits and its co-workers (University of Düsseldorf) in the course of this dissertation resulted in well diffracting crystals and illustrated that EctC forms a homodimer in solution

(Publication 5), a quaternary assembly that has also been reported for EctC from *H. elongata* in a previous study (Ono *et al*., 1999). Based on bioinformatic analysis of microbial genome databases, previous considerations of EctC suggest that the ectoine synthase is a member of the cupin family (Dunwell *et al*., 2001). Proteins of this family are built on a common structural scaffold (Agarwal *et al*., 2009; Uberto and Moomaw, 2013) allowing its pro- and eukaryotic members to perform a variety of both enzymatic and non-enzymatic functions. Strikingly, based on the data presented within this dissertation, a crystal structure of the ectoine synthase from *S. alaskensis* has been published posterior to the present dissertation demonstrating that it is indeed a member of the cupin superfamily (Widderich *et al.,* 2016d). Hence, the findings of this dissertation significantly contributed to the solution of the crystal structure of the ectoine synthase.

Collectively, the data discussed in this section significantly increase the understanding of the overall structure of the ectoine hydroxylase since they provide a detailed view into its active site and reveal the positioning and binding motifs of the various ligands within the catalytic core of this enzyme. Additionally, the successful crystallization trials of EctC provided a solid basis for the solution of the crystal structure of ectoine synthase as well. The discussed data thus (might) contribute to the decipherment of the exact catalyzed enzyme reaction mechanisms which will be discussed in the following section.

V. Reaction mechanisms of the key enzymes in ectoine biosynthesis

Based on the data discussed in the previous sections, it was yet possible to propose the reaction mechanism catalyzed by the key enzyme in hydroxyectoine biosynthesis, the ectoine hydroxylase. In terms of EctC, no structure has yet been reported that contains all ligands. Hence, a proposal of the reaction mechanism catalyzed by EctC is not possible to date.

In terms of EctD, a reaction mechanism (Fig. 11) has been proposed by Prof. Wolgang Buckel (University of Marburg) that has been described in detail in publication 1 (Widderich *et al*., 2014b) and in the book chapter (Appendix I) (Widderich *et al*., 2016c). Members of the non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily possess a mononuclear ferrous iron center that catalyzes an oxidative decarboxylation of the co-substrate 2-oxoglutarate coupled with a two-electron oxidation of the substrate (Aik *et al*., 2012; Grzyska *et al*., 2010; Hangasky *et al*., 2013; Knauer *et al*., 2012). This general theme also occurs in the catalytic core of the ectoine hydroxylase, in which the substrate ectoine is oxidized in this manner. The hydroxylation of ectoine adheres to the general catalytic mechanism of Fe(II)-and 2-oxoglutarate-dependent oxygenases (Martinez and Hausinger, 2015) and is described in detail in the following:

In the first step, the substrate-free ectoine hydroxylase is believed to contain iron(II) octahedrally coordinated by the highly conserved 2-His-1-Asp facial triad and three diffusible water molecules (Reuter *et al*., 2010). Upon binding of the co-substrate 2-oxoglutarate, two of these water molecules are thought to be released allowing subsequent binding of the substrate ectoine. In the second step, posterior to the binding of all substrates, molecular oxygen would replace the third water molecule from the mononuclear iron center implicating oxidation of iron(II) to iron(IV). The resulting highly reactive iron(IV)-peroxo species is then going to attack, in the third step, the carbonyl-C atom of the co-substrate 2-oxoglutarate with its anion, thereby yielding succinate and carbon dioxide (Riggs-Gelasco *et al*., 2004). The emerged ferryl iron(IV) = O species is then able to hydroxylate ectoine to form (*4S,5S*)-5-hydroxyectoine. Whether this occurs by insertion of the ferryl oxygen atom into the C–H bond, which would reflect a two-state reactivity mechanism, or by FeIV = O induced abstraction of the pro-*S* hydrogen at the C5-atom of ectoine and subsequent re-addition of the hydroxyl radical from iron(III)-OH to the ectoine radical intermediate with retention of configuration, which would reflect a re-bound mechanism, is unclear so far. In the final step, the product (*4S,5S*)-5-hydroxyectoine and the co-product succinate will be

released allowing three newly arriving water molecules to bind and occupy the vacant positions of the iron(II) coordination site.

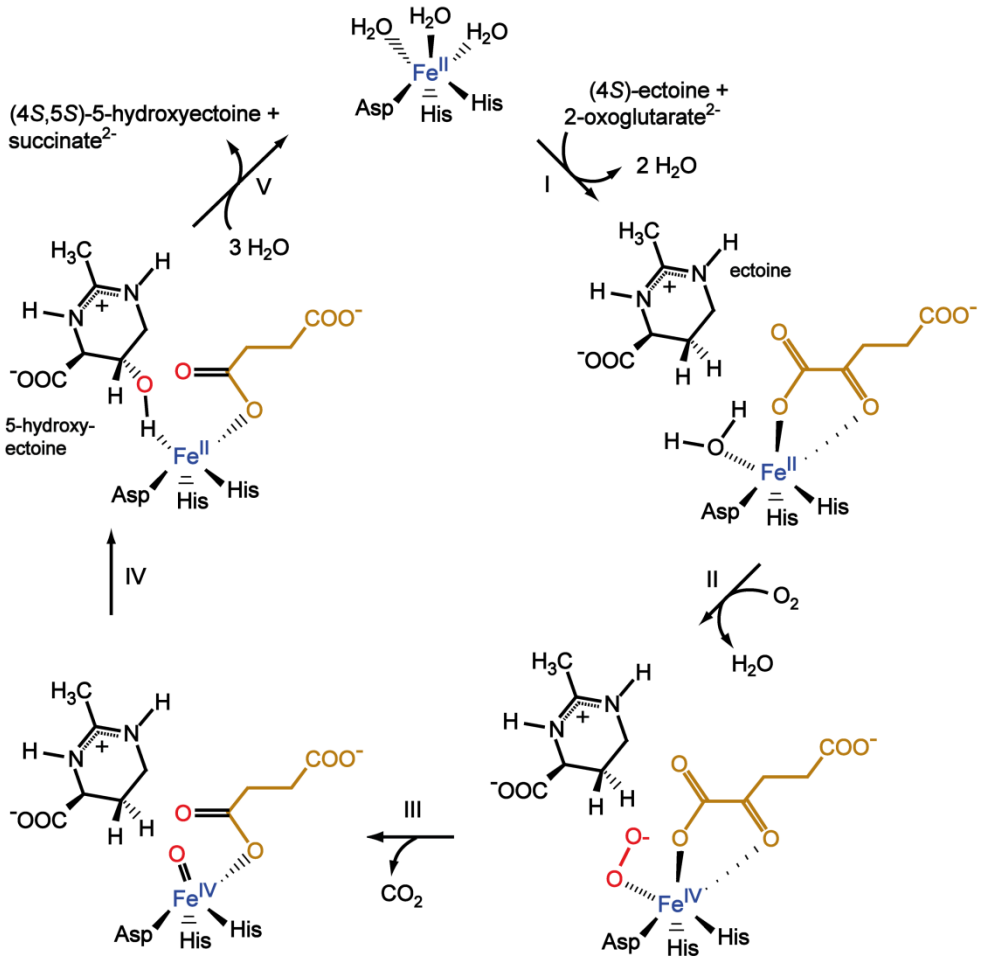


Fig 11. EctD-catalyzed reaction mechanism. For details see text. Modified from Widderich *et al.* (Widderich *et al*., 2014b) and adapted from Widderich *et al.* (Widderich *et al*., 2016c). This model was proposed by Prof. Dr. Wolfgang Buckel (see Widderich *et al.*, 2014b).

Closing remarks and perspectives

In the course of this dissertation, the phylogenetic distribution and affiliation of ectoine and hydroxyectoine biosynthetic genes was analyzed. On this basis, it was possible to further determine the salient biochemical and kinetic properties of ectoine synthases and ectoine hydroxylases from various microorganisms providing a broader analysis of their biochemistry. Finally, crystal structures of the ectoine hydroxylase in various crystal forms have been solved in collaboration with Dr. Sander Smits allowing, in connection with site-directed mutagenesis experiments, the first in-depth view into the active site of this important enzyme and the identification of amino acid residues crucial for ligand binding. These findings permitted for the first time, in collaboration with Prof. W. Buckel, a detailed proposal for its enzymatic reaction mechanism. Additionally, initial crystallization trials of the ectoine synthase resulted in well diffracting crystals that were well suited for the solution of a crystal struchture of EctC posterior to this dissertation (Widderich *et al.*, 2016d). Overall, the results obtained in the course of this dissertation provide detailed knowledge of the phylogenetic distribution, biochemistry and crystal strucuture of the key enzymes in ectoine and hydroxyectoine biosynthesis. Hence, these findings significantly contribute to the understanding of the role of ectoine and hydroxyectoine in microbial osmostress adaptation.

The findings of the present dissertation can be used as a basis to further study the ectoine synthase and other enzymes in ectoine biosynthesis, namely EctA and EctB. A crystal structure of EctA from *Bordetella parapertussis* has already been solved in complex with its ligand (PDB code: 3D3S) whereas one of EctB is still missing, and there is only a minor number of studies dealing with the biochemistry and kinetic properties of these enzymes. Hence, there are still a lot of open questions. It might thus be interesting to study these two proteins in more detail in order to unravel the biochemistry of the complete ectoine biosynthesis pathway. Moreover, analysis of other genes that have been shown in the course of this dissertation to be associated with ectoine biosynthesis could be of interest. An interesting candidate would be for example *ectR*, whose gene product has been demonstrated to mediate repression of ectoine biosynthesis in the absence of osmotic stress. However its exact underlying molecular mechanisms are still lacking, as well as a high resolution crystal structure. Another interesting approach could focus on the transport systems associated with the ectoine biosynthetic genes in order to elucidate whether they mediate the acquisition and accumulation of ectoines.

The findings of this dissertation are not only of interest in terms of basic science, they have also potential implications from a biotechnological point of view and within the framework of chemical biology. Since ectoines are found in a variety of products with medical applications, the findings of the present dissertation might contribute to improved biosynthesis of ectoines by usage of more stable enzymes or more effective producer strains and on the other hand by trials to generate chemically modified ectoine molecules. Such synthetic ectoine derivatives possessing expanded or reduced ring sizes have indeed already been reported in literature (Moritz *et al.*, 2015; Witt et *al.*, 2011). These synthetic compounds partially exhibit different properties and might thus potentially lead to interesting novel biotechnological applications and medical uses.

References

Ablinger, E., Hellweger, M., Leitgeb, S., and Zimmer, A. (2012). Evaluating the effects of buffer conditions and extremolytes on thermostability of granulocyte colony-stimulating factor using high-throughput screening combined with design of experiments. *Int J Pharm* 436: 744-752.

Agarwal, G., Rajavel, M., Gopal, B., and Srinivasan, N. (2009). Structure-based phylogeny as a diagnostic for functional characterization of proteins with a cupin fold. *PLoS ONE* 4: e5736.

Aik, W., McDonough, M.A., Thalhammer, A., Chowdhury, R., and Schofield, C.J. (2012). Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. *Curr Opin Struct Biol* 22: 691-700.

Andersson, M.M., Breccia, J.D., and Hatti-Kaul, R. (2000). Stabilizing effect of chemical additives against oxidation of lactate dehydrogenase. *Biotechnol Appl Biochem* 32 (Pt 3): 145-153.

Arakawa, T., and Timasheff, S.N. (1985). The stabilization of proteins by osmolytes. *Biophys J* 47: 411-414.

Arora, A., Ha, C., and Park, C.B. (2004). Inhibition of insulin amyloid formation by small stress molecules. *FEBS Lett* 564: 121-125.

Bashir, A., Hoffmann, T., Kempf, B., Xie, X., Smits, S.H., and Bremer, E. (2014a). Plant-derived compatible solutes proline betaine and betonicine confer enhanced osmotic and temperature stress tolerance to *Bacillus subtilis*. *Microbiology* 160: 2283-2294.

Bashir, A., Hoffmann, T., Smits, S.H., and Bremer, E. (2014b). Dimethylglycine provides salt and temperature stress protection to *Bacillus subtilis*. *Appl Environ Microbiol* 80: 2773-2785.

Bestvater, T., Louis, P., and Galinski, E.A. (2008). Heterologous ectoine production in *Escherichia coli*: by-passing the metabolic bottle-neck. *Saline Systems* 4: 12.

Blasiak, L.C., Vaillancourt, F.H., Walsh, C.T., and Drennan, C.L. (2006). Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis. *Nature* 440: 368-371.

Bolen, D.W., and Baskakov, I.V. (2001). The osmophobic effect: natural selection of a thermodynamic force in protein folding. *J Mol Biol* 310: 955-963.

Booth, I.R. (2014). Bacterial mechanosensitive channels: progress towards an understanding of their roles in cell physiology. *Curr Opin Microbiol* 18: 16-22.

Booth, I.R., and Blount, P. (2012). The MscS and MscL families of mechanosensitive channels act as microbial emergency release valves. *J Bacteriol* 194: 4802-4809.

Booth, I.R., Edwards, M.D., Black, S., Schumann, U., and Miller, S. (2007). Mechanosensitive channels in bacteria: signs of closure? *Nat Rev Microbiol* 5: 431-440.

Borges, N., Ramos, A., Raven, N.D., Sharp, R.J., and Santos, H. (2002). Comparative study of the thermostabilizing properties of mannosylglycerate and other compatible solutes on model enzymes. *Extremophiles* 6: 209-216.

Bremer, E. (2002). Adaption to changing osmolarity. In: *Bacillus subtilis and its closest relatives: from genes to cells.* Sonenshein, A.L., Hoch, J.A., and Losick, R.M. (eds). Washington D.C.: ASM Press, pp. 385-391.

Bremer, E., and Krämer, R. (2000). Coping with osmotic challenges: osmoregulation through accumulation and release of compatible solutes. In: *Bacterial Stress Responses.* Storz, G., and Hengge-Aronis, R. (eds). ASM Press., pp. 79-97.

Brill, J., Hoffmann, T., Bleisteiner, M., and Bremer, E. (2011). Osmotically controlled synthesis of the compatible solute proline is critical for cellular defense of *Bacillus subtilis* against high osmolarity. *J Bacteriol* 193: 5335-5346.

Buenger, J., and Driller, H. (2004). Ectoin: an effective natural substance to prevent UVA-induced premature photoaging. *Skin Pharmacol Physiol* 17: 232-237.

Bursy, J. (2005). Osmotisch regulierte Biosynthese der kompatiblen Solute Ectoin und Hydroxyectoin in *Salibacillus salexigens*: Biochemische Charakterisierung der Ectoin-Hydroxylase EctD und Identifizierung ihres Strukturgenes. In.: Philipps-Universität Marburg.

Bursy, J., Kuhlmann, A.U., Pittelkow, M., Hartmann, H., Jebbar, M., Pierik, A.J., and Bremer, E. (2008). Synthesis and uptake of the compatible solutes ectoine and 5-hydroxyectoine by *Streptomyces coelicolor* A3(2) in response to salt and heat stresses. *Appl Environ Microbiol* 74: 7286-7296.

Bursy, J., Pierik, A.J., Pica, N., and Bremer, E. (2007). Osmotically induced synthesis of the compatible solute hydroxyectoine is mediated by an evolutionarily conserved ectoine hydroxylase. *J Biol Chem* 282: 31147-31155.

Canovas, D., Borges, N., Vargas, C., Ventosa, A., Nieto, J.J., and Santos, H. (1999). Role of Ngamma-acetyldiaminobutyrate as an enzyme stabilizer and an intermediate in the biosynthesis of hydroxyectoine. *Appl Environ Microbiol* 65: 3774-3779.

Chastanet, A., and Losick, R. (2011). Just-in-time control of Spo0A synthesis in *Bacillus subtilis* by multiple regulatory mechanisms. *J Bacteriol* 193: 6366-6374.

Cohen, G.N., and Saint-Girons, I. (1987). Biosynthesis of threonine, lysine, and methionine. In: *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology.* Neidhardt, J.L.I.F.C., Low, K.B., Magasanik, B., Schaechter, M., and Umbarger, H.E. (eds). Washington D.C., USA: ASM Press, pp.

Coquelle, N., and Glover, J.N. (2010). FHA domain pThr binding specificity: it's all about me. *Structure* 18: 1549-1550.

Csonka, L.N. (1989). Physiological and genetic responses of bacteria to osmotic stress. *Microbiol Rev* 53: 121-147.

Csonka, L.N., and Epstein, W. (1996). Osmoregulation. In *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. Edited by F. C. Neidhard, R. Curtiss III, J. L. Ingraham *et al*., ASM Press, Washington, D. C., USA: 1210-1223.

Cyplik, P., Piotrowska-Cyplik, A., Marecik, R., Czarny, J., Drozdzynska, A., and Chrzanowski, L. (2012). Biological denitrification of brine: the effect of compatible solutes on enzyme activities and fatty acid degradation. *Biodegradation* 23: 663-672.

da Costa, M.S., Santos, H., and Galinski, E.A. (1998). An overview of the role and diversity of compatible solutes in *Bacteria* and *Archaea*. *Adv Biochem Eng Biotechnol* 61: 117-153.

Du, Y., Shi, W.W., He, Y.X., Yang, Y.H., Zhou, C.Z., and Chen, Y. (2011). Structures of the substrate-binding protein provide insights into the multiple compatible solute binding specificities of the *Bacillus subtilis* ABC transporter OpuC. *Biochem J* 436: 283-289.

Dunwell, J.M., Culham, A., Carter, C.E., Sosa-Aguirre, C.R., and Goodenough, P.W. (2001). Evolution of functional diversity in the cupin superfamily. *Trends Biochem Sci* 26: 740-746.

Dunwell, J.M., Purvis, A., and Khuri, S. (2004). Cupins: the most functionally diverse protein superfamily? *Phytochemistry* 65: 7-17.

Eichenberger, P., Fujita, M., Jensen, S.T., Conlon, E.M., Rudner, D.Z., Wang, S.T., Ferguson, C., Haga, K., Sato, T., Liu, J.S., and Losick, R. (2004). The program of gene transcription for a single differentiating cell type during sporulation in *Bacillus subtilis*. *PLoS Biol* 2: e328.

Empadinhas, N., and da Costa, M.S. (2008). Osmoadaptation mechanisms in prokaryotes: distribution of compatible solutes. *Int Microbiol* 11: 151-161.

Errington, J. (2003). Regulation of endospore formation in *Bacillus subtilis*. *Nat Rev Microbiol* 1: 117-126.

Fletcher, S.A., and Csonka, L.N. (1998). Characterization of the induction of increased thermotolerance by high osmolarity in *Salmonella*. *Int J Food Microbiol* 15: 307-317.

Francez-Charlot, A., Kaczmarczyk, A., Fischer, H.M., and Vorholt, J.A. (2015). The general stress response in *Alphaproteobacteria*. *Trends Microbiol* 23: 164-171.

Freimark, D., Sehl, C., Weber, C., Hudel, K., Czermak, P., Hofmann, N., Spindler, R., and Glasmacher, B. (2011). Systematic parameter optimization of a Me(2)SO- and serum-free cryopreservation protocol for human mesenchymal stem cells. *Cryobiology* 63: 67-75.

Galinski, E.A. (1993). Compatible solutes of halophilic eubacteria: molecular principles, water-solute interaction, stress protection. *Experientia* 49: 487-496.

Galinski, E.A. (1995). Osmoadaptation in Bacteria. *Adv Microb Physiol* 37: 273-328.

Galinski, E.A., Pfeiffer, H.P., and Trüper, H.G. (1985). 1,4,5,6-Tetrahydro-2-methyl-4-pyrimidinecarboxylic acid. A novel cyclic amino acid from halophilic phototrophic bacteria of the genus *Ectothiorhodospira*. *Eur J Biochem* 149: 135-139.

Galinski, E.A., and Trüper, H.G. (1994). Microbial behaviour in salt-stressed ecosystems. *FEMS Microbiol Rev* 15: 95-108.

Garcia-Estepa, R., Argandona, M., Reina-Bueno, M., Capote, N., Iglesias-Guerra, F., Nieto, J.J., and Vargas, C. (2006). The *ectD* gene, which is involved in the synthesis of the compatible solute hydroxyectoine, is essential for thermoprotection of the halophilic bacterium *Chromohalobacter salexigens*. *J Bacteriol* 188: 3774-3784.

Giraud, M.F., Leonard, G.A., Field, R.A., Berlind, C., and Naismith, J.H. (2000). RmlC, the third enzyme of dTDP-L-rhamnose pathway, is a new class of epimerase. *Nat Struct Biol* 7: 398-402.

Goller, K., Ofer, A., and Galinski, E.A. (1998). Construction and characterization of an NaCl-sensitive mutant of *Halomonas elongata* impaired in ectoine biosynthesis. *FEMS Microbiol Lett* 161: 293-300.

Gopal, B., Madan, L.L., Betz, S.F., and Kossiakoff, A.A. (2005). The crystal structure of a quercetin 2,3-dioxygenase from *Bacillus subtilis* suggests modulation of enzyme activity by a change in the metal ion at the active site(s). *Biochemistry* 44: 193-201.

Graf, R., Anzali, S., Buenger, J., Pfluecker, F., and Driller, H. (2008). The multifunctional role of ectoine as a natural cell protectant. *Clin Dermatol* 26: 326-333.

Grammann, K., Volke, A., and Kunte, H.J. (2002). New type of osmoregulated solute transporter identified in halophilic members of the bacteria domain: TRAP transporter TeaABC mediates uptake of ectoine and hydroxyectoine in *Halomonas elongata* DSM 2581(T). *J Bacteriol* 184: 3078-3085.

Grant, W.D. (2004). Life at low water activity. *Philos Trans R Soc Lond B Biol Sci* 359: 1249-1266; discussion 1266-1247.

Grossman, A.D. (1995). Genetic networks controlling the initiation of sporulation and the development of genetic competence in *Bacillus subtilis*. *Annu Rev Genet* 29: 477-508.

Grzyska, P.K., Appelman, E.H., Hausinger, R.P., and Proshlyakov, D.A. (2010). Insight into the mechanism of an iron dioxygenase by resolution of steps following the FeIV=HO species. *Proc Natl Acad Sci U S A* 107: 3982-3987.

Hajnal, I., Lyskowski, A., Hanefeld, U., Gruber, K., Schwab, H., and Steiner, K. (2013). Biochemical and structural characterization of a novel bacterial manganese-dependent hydroxynitrile lyase. *FEBS J* 280: 5815-5828.

Hanekop, N., Hoing, M., Sohn-Bosser, L., Jebbar, M., Schmitt, L., and Bremer, E. (2007). Crystal structure of the ligand-binding protein EhuB from *Sinorhizobium meliloti* reveals substrate recognition of the compatible solutes ectoine and hydroxyectoine. *J Mol Biol* 374: 1237-1250.

Hangasky, J.A., Taabazuing, C.Y., Valliere, M.A., and Knapp, M.J. (2013). Imposing function down a (cupin)-barrel: secondary structure and metal stereochemistry in the alphaKG-dependent oxygenases. *Metallomics* 5: 287-301.

Harishchandra, R.K., Wulff, S., Lentzen, G., Neuhaus, T., and Galla, H.J. (2010). The effect of compatible solute ectoines on the structural organization of lipid monolayer and bilayer membranes. *Biophys Chem* 150: 37-46.

Haswell, E.S., Phillips, R., and Rees, D.C. (2011). Mechanosensitive channels: what can they do and how do they do it? *Structure* 19: 1356-1369.

Hausinger, R.P. (2004). FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes. *Crit Rev Biochem Mol Biol* 39: 21-68.

Hecker, M., Pane-Farre, J., and Volker, U. (2007). SigB-dependent general stress response in *Bacillus subtilis* and related gram-positive bacteria. *Annu Rev Microbiol* 61: 215-236.

Hengge-Aronis, R. (2000). The general stress response in *Escherichia coli*. In: *Bacterial Stress Responses.* Storz, G., and Hengge-Aronis, R. (eds). Washington, D.C.: ASM Press, pp. 161-178.

Hengge-Aronis, R. (2002). Recent insights into the general stress response regulatory network in *Escherichia coli*. *J Mol Microbiol Biotechnol* 4: 341-346.

Higgins, D., and Dworkin, J. (2012). Recent progress in *Bacillus subtilis* sporulation. *FEMS Microbiol Rev* 36: 131-148.

Hoch, J.A. (1993). Regulation of the phosphorelay and the initiation of sporulation in *Bacillus subtilis*. *Annu Rev Microbiol* 47: 441-465.

Hoeppner, A., Widderich, N., Bremer, E., and Smits, S.H. (2014a). Overexpression, crystallization and preliminary X-ray crystallographic analysis of the ectoine hydroxylase from *Sphingopyxis alaskensis*. *Acta Crystallogr F Struct Biol Commun* 70: 493-496.

Hoeppner, A., Widderich, N., Lenders, M., Bremer, E., and Smits, S.H. (2014b). Crystal structure of the ectoine hydroxylase, a snapshot of the active site. *J Biol Chem* 289: 29570-29583.

Hoffmann, T., Boiangiu, C., Moses, S., and Bremer, E. (2008). Responses of *Bacillus subtilis* to hypotonic challenges: physiological contributions of mechanosensitive channels to cellular survival. *Appl Environ Microbiol* 74: 2454-2460.

Hoffmann, T., and Bremer, E. (2011). Protection of *Bacillus subtilis* against cold stress via compatible-solute acquisition. *J Bacteriol* 193: 1552-1562.

Hoffmann, T., von Blohn, C., Stanek, A., Moses, S., Barzantny, H., and Bremer, E. (2012). Synthesis, release, and recapture of compatible solute proline by osmotically stressed *Bacillus subtilis* cells. *Appl Environ Microbiol* 78: 5753-5762.

Hoffmann, T., Wensing, A., Brosius, M., Steil, L., Volker, U., and Bremer, E. (2013). Osmotic control of *opuA* expression in *Bacillus subtilis* and its modulation in response to intracellular glycine betaine and proline pools. *J Bacteriol* 195: 510-522.

Holtmann, G., and Bremer, E. (2004). Thermoprotection of *Bacillus subtilis* by exogenously provided glycine betaine and structurally related compatible solutes: involvement of Opu transporters. *J Bacteriol* 186: 1683-1693.

Holtzendorff, J., Hung, D., Brende, P., Reisenauer, A., Viollier, P.H., McAdams, H.H., and Shapiro, L. (2004). Oscillating global regulators control the genetic circuit driving a bacterial cell cycle. *Science* 304: 983-987.

Hoper, D., Volker, U., and Hecker, M. (2005). Comprehensive characterization of the contribution of individual SigB-dependent general stress genes to stress resistance of *Bacillus subtilis*. *J Bacteriol* 187: 2810-2826.

Inbar, L., Frolow, F., and Lapidot, A. (1993). The conformation of new tetrahydropyrimidine derivatives in solution and in the crystal. *Eur J Biochem* 214: 897-906.

Inbar, L., and Lapidot, A. (1988). The structure and biosynthesis of new tetrahydropyrimidine derivatives in actinomycin D producer *Streptomyces parvulus*. Use of 13C- and 15N-labeled L-glutamate and 13C and 15N NMR spectroscopy. *J Biol Chem* 263: 16014-16022.

Jebbar, M., Talibart, R., Gloux, K., Bernard, T., and Blanco, C. (1992). Osmoprotection of *Escherichia coli* by ectoine: uptake and accumulation characteristics. *J Bacteriol* 174: 5027-5035.

Jebbar, M., von Blohn, C., and Bremer, E. (1997). Ectoine functions as an osmoprtectant in *Bacillus subtilis* and is accumulated via the ABC transport system OpuC. *FEMS Microbiol Lett* 154: 325-330.

Kanapathipillai, M., Ku, S.H., Girigoswami, K., and Park, C.B. (2008). Small stress molecules inhibit aggregation and neurotoxicity of prion peptide 106-126. *Biochem Biophys Res Commun* 365: 808-813.

Kanapathipillai, M., Lentzen, G., Sierks, M., and Park, C.B. (2005). Ectoine and hydroxyectoine inhibit aggregation and neurotoxicity of Alzheimer's beta-amyloid. *FEBS Lett* 579: 4775-4780.

Kappes, R.M., Kempf, B., and Bremer, E. (1996). Three transport systems for the osmoprotectant glycine betaine operate in *Bacillus subtilis*: characterization of OpuD. *J Bacteriol* 178: 5071-5079.

Kappes, R.M., Kempf, B., Kneip, S., Boch, J., Gade, J., Meier-Wagner, J., and Bremer, E. (1999). Two evolutionarily closely related ABC transporters mediate the uptake of choline for synthesis of the osmoprotectant glycine betaine in *Bacillus subtilis*. *Mol Microbiol* 32: 203-216.

Kempf, B., and Bremer, E. (1995). OpuA, an osmotically regulated binding protein-dependent transport system for the osmoprotectant glycine betaine in *Bacillus subtilis*. *J Biol Chem* 270: 16701-16713.

Kempf, B., and Bremer, E. (1998). Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Arch Microbiol* 170: 319-330.

Khare, D., Wang, B., Gu, L., Razelun, J., Sherman, D.H., Gerwick, W.H., Hakansson, K., and Smith, J.L. (2010). Conformational switch triggered by alpha-ketoglutarate in a halogenase of curacin A biosynthesis. *Proc Natl Acad Sci U S A* 107: 14099-14104.

Kloda, A., and Martinac, B. (2002). Mechanosensitive channels of bacteria and archaea share a common ancestral origin. *Eur Biophys J* 31: 14-25.

Knauer, S.H., Hartl-Spiegelhauer, O., Schwarzinger, S., Hanzelmann, P., and Dobbek, H. (2012). The Fe(II)/alpha-ketoglutarate-dependent taurine dioxygenases from *Pseudomonas putida* and *Escherichia coli* are tetramers. *FEBS J* 279: 816-831.

Kobus, S., Widderich, N., Hoeppner, A., Bremer, E., and Smits, S.H. (2015). Overproduction, crystallization and X-ray diffraction data analysis of ectoine synthase from the cold-adapted marine bacterium *Sphingopyxis alaskensis*. *Acta Crystallogr F Struct Biol Commun* 71: 1027-1032.

Koch, A.L. (1983). The surface stress theory of microbial morphogenesis. *Adv Microb Physiol* 24: 301-366.

Kolp, S., Pietsch, M., Galinski, E.A., and Gutschow, M. (2006). Compatible solutes as protectants for zymogens against proteolysis. *Biochim Biophys Acta* 1764: 1234-1242.

Konneke, M., Bernhard, A.E., de la Torre, J.R., Walker, C.B., Waterbury, J.B., and Stahl, D.A. (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* 437: 543-546.

Konovalova, A., Petters, T., and Sogaard-Andersen, L. (2010). Extracellular biology of *Myxococcus xanthus*. *FEMS Microbiol Rev* 34: 89-106.

Korber, D.R., Choi, A., Wolfaardt, G.M., and Caldwell, D.E. (1996). Bacterial plasmolysis as a physical indicator of viability. *Appl Environ Microbiol* 62: 3939-3947.

Krulwich, T.A., Sachs, G., and Padan, E. (2011). Molecular aspects of bacterial pH sensing and homeostasis. *Nat Rev Microbiol* 9: 330-343.

Kuhlmann, A.U., and Bremer, E. (2002). Osmotically regulated synthesis of the compatible solute ectoine in *Bacillus pasteurii* and related *Bacillus* spp. *Appl Environ Microbiol* 68: 772-783.

Kuhlmann, A.U., Bursy, J., Gimpel, S., Hoffmann, T., and Bremer, E. (2008a). Synthesis of the compatible solute ectoine in *Virgibacillus pantothenticus* is triggered by high salinity and low growth temperature. *Appl Environ Microbiol* 74: 4560-4563.

Kuhlmann, A.U., Hoffmann, T., Bursy, J., Jebbar, M., and Bremer, E. (2011). Ectoine and hydroxyectoine as protectants against osmotic and cold stress: uptake through the SigB-controlled betaine-choline- carnitine transporter-type carrier EctT from *Virgibacillus pantothenticus*. *J Bacteriol* 193: 4699-4708.

Kuhlmann, S.I., Terwisscha van Scheltinga, A.C., Bienert, R., Kunte, H.J., and Ziegler, C. (2008b). 1.55 A structure of the ectoine binding protein TeaA of the osmoregulated TRAP-transporter TeaABC from *Halomonas elongata*. *Biochemistry* 47: 9475-9485.

Kundu, S. (2012). Distribution and prediction of catalytic domains in 2-oxoglutarate dependent dioxygenases. *BMC Res Notes* 5: 410.

Kung, C., Martinac, B., and Sukharev, S. (2010). Mechanosensitive channels in microbes. *Annu Rev Microbiol* 64: 313-329.

Kunst, F., and Rapoport, G. (1995). Salt stress is an environmental signal affecting degradative enzyme synthesis in *Bacillus subtilis*. *J Bacteriol* 177: 2403-2407.

Kunte, H.J., Lentzen, G., and Galinski, E.A. (2014). Industrial production of the cell protectant ectoine: protection, mechanisms, processes, and products. *Curr Biotechnol* 3: 10-25.

Kurz, M. (2008). Compatible solute influence on nucleic acids: many questions but few answers. *Saline Systems* 4: 6.

Kurz, M., Montermann, C., and Galinski, E.A. (2007). A *Pseudomonas stutzeri* surprise: production of the compatible osmolyte hydroxyectoine. *Abstr. Ann. Conf. Assoc. Gen. Appl. Microbiol.*: abstr. PT015, p. 195.

Lecher, J., Pittelkow, M., Zobel, S., Bursy, J., Bonig, T., Smits, S.H., Schmitt, L., and Bremer, E. (2009). The crystal structure of UehA in complex with ectoine-A comparison with other TRAP-T binding proteins. *J Mol Biol* 389: 58-73.

Lentzen, G., and Schwarz, T. (2006). Extremolytes: Natural compounds from extremophiles for versatile applications. *Appl Microbiol Biotechnol* 72: 623-634.

Lippert, K., and Galinski, E.A. (1992). Enzyme stabilization by ectoine-type compatible solutes: protection against heating, freezing and drying. *Appl Microbiol Biotechnol* 37: 61-65.

Lo, C.C., Bonner, C.A., Xie, G., D'Souza, M., and Jensen, R.A. (2009). Cohesion group approach for evolutionary analysis of aspartokinase, an enzyme that feeds a branched network of many biochemical pathways. *Microbiol Mol Biol Rev* 73: 594-651.

Lopez, D., and Kolter, R. (2010). Extracellular signals that define distinct and coexisting cell fates in *Bacillus subtilis*. *FEMS Microbiol Rev* 34: 134-149.

Lopez, D., Vlamakis, H., and Kolter, R. (2009). Generation of multiple cell types in *Bacillus subtilis*. *FEMS Microbiol Rev* 33: 152-163.

Louis, P., and Galinski, E.A. (1997). Characterization of genes for the biosynthesis of the compatible solute ectoine from *Marinococcus halophilus* and osmoregulated expression in *Escherichia coli*. *Microbiology* 143 (Pt 4): 1141-1149.

Lucht, J.M., and Bremer, E. (1994). Adaptation of *Escherichia coli* to high osmolarity environments: osmoregulation of the high-affinity glycine betaine transport system ProU. *FEMS Microbiol Rev* 14: 3-20.

Malin, G., and Lapidot, A. (1996). Induction of synthesis of tetrahydropyrimidine derivatives in *Streptomyces* strains and their effect on *Escherichia coli* in response to osmotic and heat stress. *J Bacteriol* 178: 385-395.

Mandic-Mulec, I., Stefanic, P., and van Elsas, J.D. (2015). Ecology of *Bacillaceae*. *Microbiol Spectr* 3: TBS-0017-2013.

Manzanera, M., Garcia de Castro, A., Tondervik, A., Rayner-Brandes, M., Strom, A.R., and Tunnacliffe, A. (2002). Hydroxyectoine is superior to trehalose for anhydrobiotic engineering of *Pseudomonas putida* KT2440. *Appl Environ Microbiol* 68: 4328-4333.

Manzanera, M., Vilchez, S., and Tunnacliffe, A. (2004). High survival and stability rates of *Escherichia coli* dried in hydroxyectoine. *FEMS Microbiol Lett* 233: 347-352.

Martinez, S., and Hausinger, R.P. (2015). Catalytic mechanisms of Fe(II)- and 2-oxoglutarate-dependent oxygenases. *J Biol Chem* 290: 20702-20711.

McDonough, M.A., Kavanagh, K.L., Butler, D., Searls, T., Oppermann, U., and Schofield, C.J. (2005). Structure of human phytanoyl-CoA 2-hydroxylase identifies molecular mechanisms of Refsum disease. *J Biol Chem* 280: 41101-41110.

Miller, K.J., and Wood, J.M. (1996). Osmoadaptation by rhizosphere bacteria. *Annu Rev Microbiol* 50: 101-136.

Molle, V., Fujita, M., Jensen, S.T., Eichenberger, P., Gonzalez-Pastor, J.E., Liu, J.S., and Losick, R. (2003). The Spo0A regulon of *Bacillus subtilis*. *Mol Microbiol* 50: 1683-1701.

Moritz, K.D., Amendt, B., Witt, E.M., and Galinski, E.A. (2015). The hydroxyectoine gene cluster of the non-halophilic acidophile *Acidiphilium cryptum*. *Extremophiles* 19: 87-99.

Moses, S., Sinner, T., Zaprasis, A., Stoveken, N., Hoffmann, T., Belitsky, B.R., Sonenshein, A.L., and Bremer, E. (2012). Proline utilization by *Bacillus subtilis*: uptake and catabolism. *J Bacteriol* 194: 745-758.

Mustakhimov, II, Reshetnikov, A.S., Fedorov, D.N., Khmelenina, V.N., and Trotsenko, Y.A. (2012). Role of EctR as transcriptional regulator of ectoine biosynthesis genes in *Methylophaga thalassica*. *Biochemistry (Mosc)* 77: 857-863.

Mustakhimov, II, Reshetnikov, A.S., Glukhov, A.S., Khmelenina, V.N., Kalyuzhnaya, M.G., and Trotsenko, Y.A. (2010). Identification and characterization of EctR1, a new transcriptional regulator of the ectoine biosynthesis genes in the halotolerant methanotroph *Methylomicrobium alcaliphilum* 20Z. *J Bacteriol* 192: 410-417.

Nannapaneni, P., Hertwig, F., Depke, M., Hecker, M., Mader, U., Volker, U., Steil, L., and van Hijum, S.A. (2012). Defining the structure of the general stress regulon of *Bacillus subtilis* using targeted microarray analysis and random forest classification. *Microbiology* 158: 696-707.

Nicolas, P., Mader, U., Dervyn, E., Rochat, T., Leduc, A., Pigeonneau, N., Bidnenko, E., Marchadier, E., Hoebeke, M., Aymerich, S., Becher, D., Bisicchia, P., Botella, E., Delumeau, O., Doherty, G., Denham, E.L., Fogg, M.J., Fromion, V., Goelzer, A., Hansen, A., Hartig, E., Harwood, C.R., Homuth, G., Jarmer, H., Jules, M., Klipp, E., Le Chat, L., Lecointe, F., Lewis, P., Liebermeister, W., March, A., Mars, R.A., Nannapaneni, P., Noone, D., Pohl, S., Rinn, B., Rugheimer, F., Sappa, P.K., Samson, F., Schaffer, M., Schwikowski, B., Steil, L., Stulke, J., Wiegert, T., Devine, K.M., Wilkinson, A.J., van Dijl, J.M., Hecker, M., Volker, U., Bessieres, P., and Noirot, P. (2012). Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science* 335: 1103-1106.

Ono, H., Sawada, K., Khunajakr, N., Tao, T., Yamamoto, M., Hiramoto, M., Shinmyo, A., Takano, M., and Murooka, Y. (1999). Characterization of biosynthetic enzymes for ectoine as a compatible solute in a moderately halophilic eubacterium, *Halomonas elongata*. *J Bacteriol* 181: 91-99.

Oren, A. (1999). Bioenergetic aspects of halophilism. *Microbiol Mol Biol Rev* 63: 334-348.

Oren, A. (2008). Microbial life at high salt concentrations: phylogenetic and metabolic diversity. *Saline Systems* 4: 2.

Pastor, J.M., Salvador, M., Argandona, M., Bernal, V., Reina-Bueno, M., Csonka, L.N., Iborra, J.L., Vargas, C., Nieto, J.J., and Canovas, M. (2010). Ectoines in cell stress protection: uses and biotechnological production. *Biotechnol Adv* 28: 782-801.

Peters, P., Galinski, E.A., and Trüper, H.G. (1990). The biosynthesis of ectoine. *FEMS Microbiol Lett* 71: 157-162.

Phillips, Z.E., and Strauch, M.A. (2002). *Bacillus subtilis* sporulation and stationary phase gene expression. *Cell Mol Life Sci* 59: 392-402.

Pittelkow, M. (2011). Synthese und physiologische Funktion der chemischen Chaperone Ectoin und Hydroxyectoin. In: Fachbereich Biologie. Marburg: Philipps-Universität Marburg.

Pittelkow, M., and Bremer, E. (2011). Cellular adjustments of *Bacillus subtilis* and other *Bacilli* to fluctuating salinities. In: *Halophiles and Hypersaline Environments: Current Research and Future Trends.* Ventosa, A., Ohren, A., and Ma, Y. (eds). Berlin: Springer, pp. 275-302.

Prabhu, J., Schauwecker, F., Grammel, N., Keller, U., and Bernhard, M. (2004). Functional expression of the ectoine hydroxylase gene (*thpD*) from *Streptomyces chrysomallus* in *Halomonas elongata*. *Appl Environ Microbiol* 70: 3130-3132.

Prescott, A.G., and Lloyd, M.D. (2000). The iron(II) and 2-oxoacid-dependent dioxygenases and their role in metabolism. *Nat Prod Rep* 17: 367-383.

Price, C.W. (2000). Protective function and regulation of the general stress response in *Bacillus subtilis* and related gram-positive bacteria. In: *Bacterial Stress Responses.* Storz, G., and Hengge-Aronis, R. (eds). Washington, D.C.: ASM Press, pp. 179-198.

Reshetnikov, A.S., Khmelenina, V.N., Mustakhimov, II, and Trotsenko, Y.A. (2011). Genes and enzymes of ectoine biosynthesis in halotolerant methanotrophs. *Methods Enzymol* 495: 15-30.

Reshetnikov, A.S., Khmelenina, V.N., and Trotsenko, Y.A. (2006). Characterization of the ectoine biosynthesis genes of haloalkalotolerant obligate methanotroph "*Methylomicrobium alcaliphilum* 20Z". *Arch Microbiol* 184: 286-297.

Reuter, K., Pittelkow, M., Bursy, J., Heine, A., Craan, T., and Bremer, E. (2010). Synthesis of 5-hydroxyectoine from ectoine: crystal structure of the non-heme iron(II) and 2-oxoglutarate-dependent dioxygenase EctD. *PLoS ONE* 5 (5): e10647. doi:10610.11371/journal.pone.0010647.

Rhodes, M.E., Fitz-Gibbon, S.T., Oren, A., and House, C.H. (2010). Amino acid signatures of salinity on an environmental scale with a focus on the Dead Sea. *Environ Microbiol* 12: 2613-2623.

Riggs-Gelasco, P.J., Price, J.C., Guyer, R.B., Brehm, J.H., Barr, E.W., Bollinger, J.M., Jr., and Krebs, C. (2004). EXAFS spectroscopic evidence for an Fe=O unit in the Fe(IV) intermediate observed during oxygen activation by taurine:alpha-ketoglutarate dioxygenase. *J Am Chem Soc* 126: 8108-8109.

Roychoudhury, A., Haussinger, D., and Oesterhelt, F. (2012). Effect of the compatible solute ectoine on the stability of the membrane proteins. *Protein Pept Lett* 19: 791-794.

Ruzal, S.M., Lopez, C., Rivas, E., and Sanchez-Rivas, C. (1998). Osmotic strength blocks sporulation at stage II by impeding activation of early sigma factors in *Bacillus subtilis*. *Curr Microbiol* 36: 75-79.

Ruzal, S.M., and Sanchez-Rivas, C. (1998). In *Bacillus subtilis* DegU-P is a positive regulator of the osmotic response. *Curr Microbiol* 37: 368-372.

Sauer, T., and Galinski, E.A. (1998). Bacterial milking: A novel bioprocess for production of compatible solutes. *Biotechnol Bioeng* 59: 128.

Schiraldi, C., Maresca, C., Catapano, A., Galinski, E.A., and De Rosa, M. (2006). High-yield cultivation of *Marinococcus* M52 for production and recovery of hydroxyectoine. *Res Microbiol* 157: 693-699.

Schnoor, M., Voß, P., Cullen, P., Böking, T., Galla, H.-J., Galinski, E.A., and Lorkowski, S. (2004). Characterization of the synthetic compatible solute homoectoine as a potent PCR enhancer. *Biochem Biophys Res Commun* 322: 867-872.

Schwibbert, K., Marin-Sanguino, A., Bagyan, I., Heidrich, G., Lentzen, G., Seitz, H., Rampp, M., Schuster, S.C., Klenk, H.P., Pfeiffer, F., Oesterhelt, D., and Kunte, H.J. (2011). A blueprint of ectoine metabolism from the genome of the industrial producer *Halomonas elongata* DSM 2581 T. *Environ Microbiol* 13: 1973-1994.

Setlow, P. (2014). Germination of spores of *Bacillus* species: what we know and do not know. *J Bacteriol* 196: 1297-1305.

Silvennoinen, L., Sandalova, T., and Schneider, G. (2009). The polyketide cyclase RemF from *Streptomyces resistomycificus* contains an unusual octahedral zinc binding site. *FEBS Lett* 583: 2917-2921.

Sinai, L., Rosenberg, A., Smith, Y., Segev, E., and Ben-Yehuda, S. (2015). The molecular timeline of a reviving bacterial spore. *Mol Cell* 57: 695-707.

Smirnoff, N., and Cumbes, Q.J. (1989). Hydroxyl radical scavenging activity of compatible solutes. *Phytochem* 28: 1057-1060.

Sonenshein, A.L. (2000). Control of sporulation initiation in *Bacillus subtilis*. *Curr Opin Microbiol* 3: 561-566.

Spiegelhalter, F., and Bremer, E. (1998). Osmoregulation of the *opuE* proline transport gene from *Bacillus subtilis*: contributions of the Sigma A- and Sigma B-dependent stress-responsive promoters. *Mol Microbiol* 29: 285-296.

Steil, L., Serrano, M., Henriques, A.O., and Volker, U. (2005). Genome-wide analysis of temporally regulated and compartment-specific gene expression in sporulating cells of *Bacillus subtilis*. *Microbiology* 151: 399-420.

Storz, G., and Hengge-Aronis, R. (2000). *Bacterial Stress Responses*. ASM Press, Washington, D.C.

Stöveken, N., Pittelkow, M., Sinner, T., Jensen, R.A., Heider, J., and Bremer, E. (2011). A specialized aspartokinase enhances the biosynthesis of the osmoprotectants ectoine and hydroxyectoine in *Pseudomonas stutzeri* A1501. *J Bacteriol* 193: 4456-4468.

Straganz, G.D., and Nidetzky, B. (2006). Variations of the 2-His-1-carboxylate theme in mononuclear non-heme FeII oxygenases. *Chembiochem* 7: 1536-1548.

Stragier, P., and Losick, R. (1996). Molecular genetics of sporulation in *Bacillus subtilis*. *Annu Rev Genet* 30: 297-241.

Sturm, A., and Dworkin, J. (2015). Phenotypic diversity as a mechanism to exit cellular dormancy. *Curr Biol* 25: 2272-2277.

Sydlik, U., Gallitz, I., Albrecht, C., Abel, J., Krutmann, J., and Unfried, K. (2009). The compatible solute ectoine protects against nanoparticle-induced neutrophilic lung inflammation. *Am J Respir Crit Care Med* 180: 29-35.

Tanne, C., Golovina, E.A., Hoekstra, F.A., Meffert, A., and Galinski, E.A. (2014). Glass-forming property of hydroxyectoine is the cause of its superior function as a desiccation protectant. *Front Microbiol* 5: 150.

Tatzelt, J., Prusiner, S.B., and Welch, W.J. (1996). Chemical chaperones interfere with the formation of scrapie prion protein. *Embo J* 15: 6363-6373.

Uberto, R., and Moomaw, E.W. (2013). Protein similarity networks reveal relationships among sequence, structure, and function within the cupin superfamily. *PLoS ONE* 8: e74477.

Vargas, C., Argandona, M., Reina-Bueno, M., Rodriguez-Moya, J., Fernandez-Aunion, C., and Nieto, J.J. (2008). Unravelling the adaptation responses to osmotic and temperature stress in *Chromohalobacter salexigens*, a bacterium with broad salinity tolerance. *Saline Systems* 4: 14.

Ventosa, A., Nieto, J.J., and Oren, A. (1998). Biology of moderately halophilic aerobic bacteria. *Microbiol Mol Biol Rev* 62: 504-544.

von Blohn, C., Kempf, B., Kappes, R.M., and Bremer, E. (1997). Osmostress response in *Bacillus subtilis*: characterization of a proline uptake system (OpuE) regulated by high osmolarity and the alternative transcription factor Sigma B. *Mol Microbiol* 25: 175-187.

Wei, L., Wedeking, A., Buttner, R., Kalff, J.C., Tolba, R.H., and van Echten-Deckert, G. (2009). A natural tetrahydropyrimidine protects small bowel from cold ischemia and subsequent warm in vitro reperfusion injury. *Pathobiology* 76: 212-220.

Widderich, N. (2012). Biochemische Charakterisierung von Ectoin-Hydroxylasen aus extremophilen Mikroorganismen. In: Biology. Philipps-Universität Marburg.

Widderich, N., Bremer, E., and Smits, S.H.J. (2016c). The ectoine hydroxylase: a non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase. In: *Encyclopedia of Inorganic and Bioinorganic Chemistry.* Messerschmidt, A. (ed). Wiley Online Library.

Widderich, N., Czech, L., Elling, F.J., Konneke, M., Stöveken, N., Pittelkow, M., Riclea, R., Dickschat, J.S., Heider, J., and Bremer, E. (2016a). Strangers in the archaeal world: osmostress-responsive biosynthesis of ectoine and hydroxyectoine by the marine thaumarchaeon *Nitrosopumilus maritimus*. *Environ Microbiol*. 18: 1227-1248.

Widderich, N., Hoeppner, A., Pittelkow, M., Heider, J., Smits, S.H., and Bremer, E. (2014a). Biochemical properties of ectoine hydroxylases from extremophiles and their wider taxonomic distribution among microorganisms. *PLoS One* 9: e93809.

Widderich, N., Pittelkow, M., Hoeppner, A., Mulnaes, D., Buckel, W., Gohlke, H., Smits, S.H., and Bremer, E. (2014b). Molecular dynamics simulations and structure-guided mutagenesis provide insight into the architecture of the catalytic core of the ectoine hydroxylase. *J Mol Biol* 426: 586-600.

Widderich, N., Rodrigues, C.D., Commichau, F.M., Fischer, K.E., Ramirez-Guadiana, F.H., Rudner, D.Z., and Bremer, E. (2016b). Salt-sensitivity of SigH and Spo0A prevents

sporulation of *Bacillus subtilis* at high osmolarity avoiding death during cellular differentiation. *Mol Microbiol*. 100, 108-124.

Widderich, N., Kobus, S., Hoeppner, A., Riclea, R., Seubert, A., Dickschat, J.S., Heider, J., Smits, S.H.J., and Bremer, E. (2016d). Biochemistry and crystal structure of ectoine synthase: a metal-containing member of the cupin superfamily. *PLoS One* 11: e0151285.

Wilson, M.E., Maksaev, G., and Haswell, E.S. (2013). MscS-like mechanosensitive channels in plants and microbes. *Biochemistry* 52: 5708-5722.

Witt, E. (2012). Nebenreaktionen der Ectoin-Synthase aus *Halomonas elongata* DSM 2581^T und Entwicklung eines salzinduzierten Expressionssystems. In: Biotechnologie. Rheinische Friedrich-Wilhelms-Universität Bonn.

Witt, E.M., Davies, N.W., and Galinski, E.A. (2011). Unexpected property of ectoine synthase and its application for synthesis of the engineered compatible solute ADPC. *Appl Microbiol Biotechnol* 91: 113-122.

Wong, C., Fujimori, D.G., Walsh, C.T., and Drennan, C.L. (2009). Structural analysis of an open active site conformation of nonheme iron halogenase CytC3. *J Am Chem Soc* 131: 4872-4879.

Wong, S.D., Srnec, M., Matthews, M.L., Liu, L.V., Kwak, Y., Park, K., Bell, C.B., 3rd, Alp, E.E., Zhao, J., Yoda, Y., Kitao, S., Seto, M., Krebs, C., Bollinger, J.M., Jr., and Solomon, E.I. (2013). Elucidation of the Fe(IV)=O intermediate in the catalytic cycle of the halogenase SyrB2. *Nature* 499: 320-323.

Wood, J.M. (2011). Bacterial osmoregulation: a paradigm for the study of cellular homeostasis. *Annu Rev Microbiol* 65: 215-238.

Wood, J.M., Bremer, E., Csonka, L.N., Kraemer, R., Poolman, B., van der Heide, T., and Smith, L.T. (2001). Osmosensing and osmoregulatory compatible solute accumulation by bacteria. *Comp Biochem Physiol A Mol Integr Physiol* 130: 437-460.

Xu, Q., Grant, J., Chiu, H.J., Farr, C.L., Jaroszewski, L., Knuth, M.W., Miller, M.D., Lesley, S.A., Godzik, A., Elsliger, M.A., Deacon, A.M., and Wilson, I.A. (2014). Crystal structure of a member of a novel family of dioxygenases (PF10014) reveals a conserved cupin fold and active site. *Proteins* 82: 164-170.

Young, J.W., Locke, J.C., and Elowitz, M.B. (2013). Rate of environmental change determines stress response specificity. *Proc Natl Acad Sci U S A* 110: 4140-4145.

Zhang, Z., Kochan, G.T., Ng, S.S., Kavanagh, K.L., Oppermann, U., Schofield, C.J., and McDonough, M.A. (2011). Crystal structure of PHYHD1A, a 2OG oxygenase related to phytanoyl-CoA hydroxylase. *Biochem Biophys Res Commun* 408: 553-558.

Appendix

I. Review Article/Book chapter

A book chapter entitled ‘The ectoine hydroxylase: a non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase’ has also been written within the course of this dissertation.

This chapter has been published in the ‘Encyclopedia of Inorganic and Bioinorganic Chemistry’ und the following DOI: 10.1002/9781119951438.eibc2440.

Widderich, N., Bremer, E. & Smits, S.H.J. The ectoine hydroxylase: a non-heme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase. In: *Encyclopedia of Inorganic and Bioinorganic Chemistry*. Messerschmidt, A. (ed). Wiley Online Library. (2016). doi: 10.1002/9781119951438.eibc2440.

This review article summarizes and compares the various solved crystal structures of the ectoine hydroxylase. It is mainly focused on the iron-binding site of EctD and outlines the importance of the iron co-factor for catalysis and enzyme functionality.

Personal contribution:
I prepared Figures and Tables (together with Sander Smits) and contributed to the writing of the book chapter together with my supervisor Erhard Bremer as well as Sander Smits.

The ectoine hydroxylase: a nonheme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase

Nils Widderich[†,‡], *Erhard Bremer*[†,‡], *and Sander HJ Smits*[*]

[†]Department of Biology, Laboratory for Microbiology, Philipps-University Marburg, Marburg, Germany
[‡]LOEWE-Center for Synthetic Microbiology, Philipps-University Marburg, Marburg, Germany
[*]Institute for Biochemistry, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

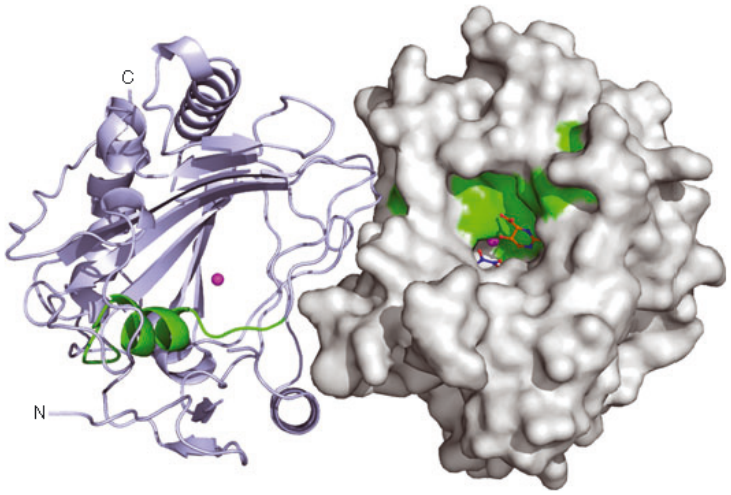
FUNCTIONAL CLASS

Enzyme; nonheme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily (EC 1.14.11). Members of this superfamily are versatile catalysts whose enzyme reactions include hydroxylations, dimethlyations, desaturations, epimerizations, cyclizations, halogenations, and ring expansions. Described here is the ectoine hydroxylase (EctD) that converts ectoine to 5-hydroxyectoine in an iron(II)-dependent manner.

The enzyme EctD hydroxylates ectoine *via* a region-selective and stereospecific enzymatic reaction forming the stress-protectant and chemical chaperone 5-hydroxyectoine [(4*S*,5*S*)-2-methyl-5-hydroxy-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid].[1,2]

OCCURRENCE

Synthesis of ectoine [(*S*)-2-methyl-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid] ensues from L-aspartate-β-semialdehyde, a central intermediate in microbial amino acid metabolism and cell wall synthesis, and comprises the sequential activities of the L-2,4-diaminobutyrate transaminase (EctB; EC 2.6.1.76), 2,4-diaminobutyrate acetyltransferase (EctA; EC 2.3.1.178), and ectoine synthase (EctC; EC 4.2.1.108). The ectoine derivative



3D Structure Structure of the dimeric ectoine hydroxylase from *S. alaskensis* (*Sa*EctD). One monomer is represented as cartoon representation. The other monomer is highlighted as a surface with the bound 2-oxoglutarate substrate (shown as blue sticks) and the 5-hydroxyectoine product (shown in orange). The bound Fe(II) ligand is shown as spheres colored in magenta (PDB code: 4Q5O). The signature sequence motif of *Sa*EctD is colored in green. These figures were prepared using PyMOL (www.pymol.org).

Encyclopedia of Inorganic and Bioinorganic Chemistry, Online © 2011–2016 John Wiley & Sons, Ltd.

This article was published in the *Encyclopedia of Inorganic and Bioinorganic Chemistry* in 2016 by John Wiley & Sons, Ltd.
DOI: 10.1002/9781119951438.eibc2440

The ectoine hydroxylase: a nonheme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase

5-hydroxyectoine [(4*S*,5*S*)-5-hydroxy-2-methyl-1,4,5,6-tetrahydropyrimidine-4-carboxylic acid] is formed only by a subgroup of the ectoine producers and is catalyzed by the ectoine hydroxylase (EctD) (EC 1.14.11). The expression of the ectoine biosynthetic genes (*ectABC*) is typically osmotically induced and this operon might also comprise the structural gene (*ectD*) for the ectoine hydroxylase. However, *ectD* genes are also often found separated from the *ectABCD* gene cluster.[1,3,4] In a recent survey of putative hydroxyectoine-producing microorganisms, 72% possess *ectD* genes that are located next to the *ectABC* gene cluster. The highest number of *ectD* genes are observed in the actinobacterial ectoine producers, in particular, in strains of the genera *Streptomyces*, *Actinomadura*, fast-growing species of *Mycobacterium*, in most species of the orders *Pseudonocardiales*, *Glycomycetales*, the genus *Nocardiopsis*, and the phylogenetically basal genus *Nitriliruptor*. There are only eight archaeal representatives predicted to synthesize 5-hydroxyectoine. Each of these is a member of the aerobic thaumarchaeal genus *Nitrosopumilus*[5] and they share very similar *ectD* gene products with those of the gamma proteobacterial genus *Nitrosococcus*. As the ectoine hydroxylase is oxygen dependent,[1,6] it is not surprising that no *ectD* gene is found in any of the anaerobic ectoine-producing bacterial or archaeal microorganisms. The EctD protein is frequently misannotated in genome sequences as proline hydroxylases or phytanoyl hydroxylases that, similar to the ectoine hydroxylase, also belong to the nonheme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily.[7–9] *Bona fide* EctD-type proteins can, however, easily be distinguished from other members of the nonheme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily by the presence of a highly conserved signature sequence motif.[1,10] This 17-amino acid region is highly conserved in the all analyzed ectoine hydroxylases, contains amino acids functionally important for iron and substrate bindings, and is an important determinant for the overall structure of the ectoine hydroxylase (see below).[6,11]

BIOLOGICAL FUNCTION

Ectoine (Figure 1) is a well-recognized compatible solute and microbial stress protectant.[12,13] It is synthesized widely by members of the *Bacteria* and a few *Archaea*[5,11,13,14] as an adaptive response to osmotic stress[1,14–20] and growth-restricting extremes in high and low temperatures.[4,17] Some ectoine producers[11] also synthesize a derivative of ectoine, in which a hydroxy group is stereospecifically attached to the pro-*S* hydrogen at position C5 in the tetrahydropyrimidine ring to yield 5-hydroxyectoine (Figure 1). This reaction is carried out by the ectoine hydroxylase EctD. 5-hydroxyectoine is often superior to its precursor molecule ectoine in protecting microorganisms against environmentally imposed stresses.[4,18]

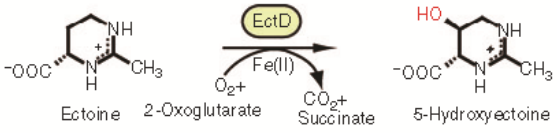


Figure 1 Scheme of the EctD-catalyzed reaction. The iron(II) containing ectoine hydroxylase converts ectoine, molecular oxygen, and 2-oxoglutarate into 5-hydroxyectoine, CO_2, and succinate.

AMINO ACID SEQUENCE INFORMATION

The following nine EctD proteins have been biochemically characterized (accession codes are derived from the UniProt database). Eight are derived from members of the *Bacteria* and one is derived from the Thaumarchaeon *N. maritimus*:

- *Halomonas elongata: He*EctD; 332 amino acids; p*I*: 5.8, protein accession code: CBV43892.
- *Acidiphilium cryptum: Ac*EctD; 306 amino acids; p*I*: 5.8, protein accession code: ABQ32201.
- *Alkalilimnicola ehrlichii: Ae*EctD; 302 amino acids; p*I*: 5.7, protein accession code: AER00257.
- *Sphingopyxis alaskensis: Sa*EctD; 306 amino acids; p*I*: 5.5, protein accession code: ABF54657.
- *Paenibacillus lautus: Pl*EctD; 302 amino acids; p*I*: 5.6, protein accession code: ACX67869.
- *Pseudomonas stutzeri: Ps*EctD; 302 amino acids; p*I*: 5.5, protein accession code: CBM40642.
- *Virgibacillus salexigens: Vs*EctD; 300 amino acids; p*I*: 5.8, protein accession code: AY935522.
- *Streptomyces coelicolor: Sc*EctD; 299 amino acids; p*I*: 5.1, protein accession code: Q93RV9.
- *Nitrosopumilus maritimus: Nm*EctD; 304 amino acids; p*I*: 5.8, protein accession code: ABX13239.

These nine EctD proteins display a degree of amino acid sequence identity between 66% and 35% toward each other. Amino acids involved in binding the Fe(II) cofactor, the cosubstrate 2-oxoglutarate, and the substrate ectoine are highly conserved in the large EctD protein family.[11]

PROTEIN PRODUCTION AND PURIFICATION

So far, nine ectoine hydroxylases have been characterized biochemically.[1,5,11] To overexpress these proteins, the corresponding *ectD* genes were either amplified from genomic DNA of the corresponding microorganisms or synthesized as codon-optimized versions for their expression in *Escherichia coli*. They were cloned into the pASK-IBA3 overexpression vector (IBA GmbH, Göttingen, Germany) and placed under the transcriptional control

This article was published in the *Encyclopedia of Inorganic and Bioinorganic Chemistry* in 2016 by John Wiley & Sons, Ltd.
DOI: 10.1002/9781119951438.eibc2440

of the TetR-responsive and anhydrotetracycline (AHT)-inducible *tet*-promoter. The resulting EctD proteins contain a *Strep*-tag affinity peptide at their carboxy terminal end to allow purification by affinity chromatography. EctD proteins have been successfully overproduced in *E. coli* BL21 cells. For these experiments, cultures were grown at 37 °C and 180 rpm until they reached an OD_{578} of 0.5; subsequently, the temperature and the shaking speed were reduced to 30 °C and 100 rpm, respectively. At OD_{578} of 0.7, overexpression of the *ectD* genes was induced by the addition of AHT. After 2 h of further growth, cells were harvested and pellets were frozen and stored at −80 °C until further usage.

For purification of the EctD proteins, the pellets were resuspended in 20 mM 2-[[1,3-dihydroxy-2-(hydroxymethyl)propan-2-yl]amino]ethanesulfonic acid (TES), pH 8, 100 mM KCl and were then disrupted by passing them three times through a French Pressure Cell Press. Cellular debris was removed by ultracentrifugation and the cleared supernatant was loaded onto a Strep-Tactin Superflow column that had been equilibrated with five bed volumes of the resuspension buffer (20 mM TES, pH 8, 100 mM KCl). The column was then washed with 10 column volumes of resuspension buffer. The EctD proteins were eluted from the affinity chromatography material with three column volumes of resuspension buffer that contained 2.5 mM desthiobiotin. The eluted EctD proteins were concentrated with Vivaspin 6 columns to a final concentration of about 5–10 mg ml^{-1}. Two hundred to three hundred milligrams EctD protein was routinely obtained through this overproduction and purification scheme per liter of cell culture for the codon-optimized genes, whereas nonoptimized version obtained about 50–100 mg. The protein concentration was measured using the Pierce™ BCA Protein Assay Kit (Thermo Scientific, Schwerte, Germany) and the specific extinction coefficient at 280 nm and the specific molecular mass of the full length EctD proteins including the *Strep*-tagII affinity peptide (Table 1).

ACTIVITY ASSAY

The enzymatic activity of the EctD proteins and their mutants was assayed in 30 µl reaction volumes that typically contained 10 mM TES buffer (pH 7.5), 10 mM 2-oxoglutarate, 6 mM ectoine, 1 mM $FeSO_4$, and various amounts of the purified EctD protein. Once the optimal conditions with respect to temperature, pH value, and salt concentration had been determined, the kinetic parameters the EctD enzyme for its cosubstrate 2-oxoglutarate, and its substrate ectoine were determined. The enzyme reaction mixtures were incubated at 32 °C for 20 min in a thermomixer (Eppendorf, Hamburg, Germany) with vigorous shaking providing enough oxygen for the O_2-dependent EctD enzyme. The enzyme reactions were stopped by the addition of 30 µl of 100% acetonitrile to the reaction mixtures and the samples were then immediately centrifuged (10 min, 4 °C, 32 000*g*) to remove the denatured EctD protein. Formation of 5-hydroxyectoine from the substrate ectoine was determined by loading 20 µl of the supernatant of the cleared enzyme reaction mixture onto a GROM-SIL 100 Amino-1PR column (125 mm by 4 mm; 3-µm particle size) (GROM, Rottenburg-Hailfingen, Germany) attached to a UV–visible detector system (LINEAR UVIS 205; SYKAM, Fürstenfeldbruck, Germany) and using 80% acetonitrile as the running solvent. Ectoine and 5-hydroxyectoine can readily be separated from each other by this chromatographic procedure (Figure 2). By observing their absorbance at 210 nm,[1,16] the amounts of the remaining ectoine and that of the newly formed 5-hydroxyectoine can be quantitated with the ChromStar 7.0 software package (SYKAM, Fürstenfeldbruck, Germany). Known quantities of commercial available ectoine and hydroxyectoine reference compounds (http://www.bitop.de/cms/website.php?id=/de/index/extremolyte.htm) were used as benchmarks to assess the catalytic efficiency of the various studied EctD enzymes.

MOLECULAR CHARACTERIZATION

Most of the studied ectoine hydroxylases have been characterized biochemically and kinetically (Table 1) in a similar

Table 1 Kinetic parameters of investigated ectoine hydroxylases[5,11,18]

EctD from	Length (AA)	MW (kDa)	p*I*	K_m (mM ectoine)	V_{max} (U/mg)	k_{cat} (s^{-1})	K_m (mM 2-oxoglutarate)
V. salexigens	300	34.4	5.8	5.9 ± 0.3	6.4 ± 0.2	3.2	4.9 ± 0.3
S. alaskensis	306	34.1	5.5	9.8 ± 0.5	1.0 ± 0.2	1.7	2.7 ± 0.3
H. elongata	332	37.4	5.8	5.7 ± 0.6	2.5 ± 0.2	1.0	4.8 ± 0.4
P. stutzeri	302	34.2	5.5	6.2 ± 0.4	6.7 ± 0.2	3.6	4.6 ± 0.5
P. lautus	302	34.8	5.6	9.5 ± 0.7	1.3 ± 0.1	0.6	3.9 ± 0.2
A. ehrlichii	302	34.3	5.7	9.0 ± 0.3	1.0 ± 0.1	0.7	5.0 ± 0.3
A. cryptum	306	34.1	5.8	10.0 ± 0.6	2.8 ± 0.3	2.3	4.1 ± 0.4
N. maritimus	304	34.7	5.8	3.8 ± 0.5	1.8 ± 0.1	1.0	3.1 ± 0.6
S. coelicolor	299	32.9	5.1	2.6 ± 0.2	20 ± 1	–	6.2 ± 0.2

This article was published in the *Encyclopedia of Inorganic and Bioinorganic Chemistry* in 2016 by John Wiley & Sons, Ltd.
DOI: 10.1002/9781119951438.eibc2440

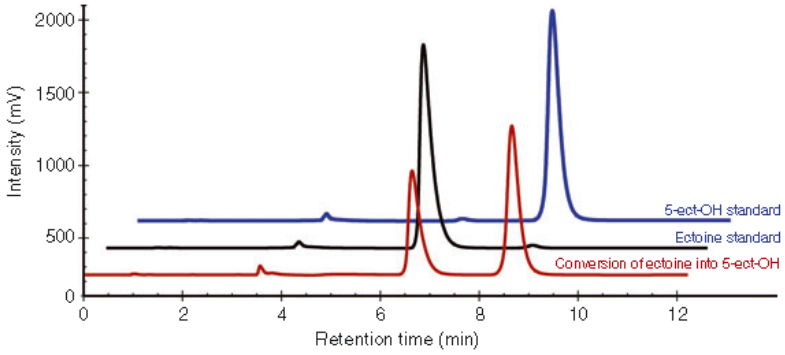


Figure 2 Representative HPLC chromatograms of ectoine and 5-hydroxyectoine. Standard solutions of ectoine (black) and 5-hydroxyectoine (blue) are depicted and a representative chromatogram of an *in vitro* conversion of ectoine into 5-hydroxyectoine (red) by the purified *Sa*EctD enzyme is shown.

way, allowing a direct comparison of enzymes originating from microorganisms living in rather different habitats and members from both the bacterial and archaeal kingdoms.[1,5,11] The temperature and pH optima ranged between 30 and 40 °C and 7.0 and 8.0, respectively, and each of the enzymes could be activated by the addition of salts. With respect to their kinetic parameters, their affinity (K_m) values for ectoine ranged between 5.5 and 10.0 mM and their maximal activity (V_{max}) between 1.0 and 7.0 U/mg of protein (Table 1). Despite the fact that the original host organisms for the studied EctD proteins belong to different bacterial phyla and inhabit diverse ecological niches, the biochemical and kinetic properties of the different studied ectoine hydroxylases are very similar.[1,5,11] Hence, the properties of the so far studied nine ectoine hydroxylases can probably be regarded as representative for the entire extended EctD protein family.[5,11]

METAL CONTENT

Biochemical and structural assessment of the ectoine hydroxylases revealed that the enzyme belongs to the nonheme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase superfamily.[1,10,21] Hence, the EctD enzyme possesses iron(II) as its cofactor. Presence of iron in EctD protein preparations was determined using a photometric assay described by Lovenberg *et al.*[22] In this assay, 10 nmol of the proteins were heated at 80 °C in 1% HCl for denaturation of the protein. The iron content of the samples was then determined by the addition of phenanthroline, which, in complex with iron, adopts a purple/red color that can be measured photometrically at 512 nm. Performing such an assay with the different, in *E. coli* recombinantly produced EctD proteins, resulted in an iron content of roughly 85–95%.

X-RAY STRUCTURES

Crystallization of VsEctD and SaEctD

The EctD enzyme has been crystallized from two different microorganisms, the cold-adapted *Sphingopyxis alaskensis* (*Sa*EctD) and the salt-tolerant *Virgibacillus salexigens* (VsEctD). The *Sa*EctD protein was quite challenging to crystallize, and most of the crystals obtained in the initial screen were of low quality. To obtain better quality crystals, the detergent screens from Hampton Research, Aliso Viejo, USA, were used as described.[23] After several rounds of optimization and buffer adjustments, crystals for X-ray crystallography were obtained, which finally resulted in structures of *Sa*EctD in the apo- and iron-bound form. The crystallization procedures are described in detail below.

Apo-*Sa*EctD: crystallization trials were performed using the sitting-drop vapor diffusion method at 20 °C. 1.5 µl of the homogenous protein solution of *Sa*EctD (10 mg ml^{-1} in 20 mM TES, pH 7.5, 80 mM NaCl) was mixed with 1.5 µl reservoir solution containing 100 mM 2-(N-morpholino)ethanesulfonic acid (MES), pH 6.0, 200 mM Ca-acetate, 27–32% (w/v) polyethyleneglycol (PEG) 400, and 1.5 mM *n*-dodecyl-*N*,*N*-dimethylglycine and equilibrated over 300 µl reservoir solution. Crystals grew within 6–12 days to their final size of around 30 µm × 30 µm × 50 µm. This resulted in crystals exhibiting a C222 symmetry with a unit cell of $a = 83.48$ Å, $b = 86.51$ Å, $c = 95.34$ Å, and $\alpha = \beta = \gamma = 90°$. These crystals diffracted to a resolution of 2.1 Å and one monomer of the apo-*Sa*EctD appeared to be present in the asymmetric unit (ASU) of the crystal.

Crystallization of the *Sa*EctD protein in complex with its iron catalyst was carried out as described above for apo-*Sa*EctD protein except that the protein solution was premixed with Fe(II)Cl_2 (final concentration of 4 mM) and equilibrated for 10–15 min on ice. In addition, the solution contained 3.5 mM n-dodecyl-N,N-dimethylglycine.

This article was published in the *Encyclopedia of Inorganic and Bioinorganic Chemistry* in 2016 by John Wiley & Sons, Ltd.
DOI: 10.1002/9781119951438.eibc2440

Crystals grew within 6–12 days at 20 °C to their final size of around 40 µm × 40 µm × 180 µm. These types of crystals diffracted to 2.6 Å resolution and exhibited an orthorhombic $P2_12_12_1$ symmetry with a unit cell of $a = 78.16$ Å, $b = 87.52$ Å, $c = 96.05$ Å, and $\alpha = \beta = \gamma = 90°$. Two monomers of *Sa*EctD were present in the ASU of the crystal. The crystal contact was mediated by n-dodecyl-N,N-dimethylglycine, a compound that had been added to the *Sa*EctD protein before crystallization.

To obtain a view of the ligand binding within the *Sa*EctD protein, crystallization was performed with the cosubstrate 2-oxoglutarate and the reaction product 5-hydroxyectoine. Using this combination, it was believed that the *Sa*EctD protein would be arrested in a ligand bound state, which then can be used for crystallization. Unfortunately, the conditions found for the apo- and Fe-bound *Sa*EctD structure did not yield crystals diffracting to high resolution. Furthermore, the structure calculated from the collected datasets (at a maximum of 4.2–4.5 Å) revealed that none of the two ligands was present within the crystal.

To achieve this, another intensive round of screening was necessary to obtain crystals of *Sa*EctD in complex with Fe, 2-oxoglutarate, and 5-hydroxyectoine. Finally, this succeeded by adding the substrates sequentially, as done before to obtain crystals of other protein–ligand complexes.[24] The *Sa*EctD protein was mixed with Fe(II)Cl_2 as described and subsequently 2-oxoglutarate was added to the *Sa*EctD to a final concentration of 40 mM. After 30 min of incubation on ice, the third compound, 5-hydroxyectoine was added to the solution at a final concentration of 40 mM. This mixture was then incubated for 1 h on ice before crystallization trials were conducted. The order in which the ligands were added to the protein solution appeared to be crucial. Substrate-bound *Sa*EctD crystals were grown by mixing 1.5 µl protein solution with 1.5 µl reservoir containing 100 mM MES, pH 6.0, 200 mM Ca-acetate, 30% (w/v) PEG 400, and 25 mM *n*-Octyl-β-D-glycoside. After 6–12 days, they reached their final size of around 40 µm × 40 µm × 70 µm. All crystals were cryoprotected by carefully adding 1 µl of 100% glycerol to the crystallization drop before the crystals were frozen in liquid nitrogen. This crystal diffracted to 2.6 Å resolution and exhibited an orthorhombic $P2_12_12_1$ symmetry with a unit cell of $a = 81.00$ Å, $b = 87.08$ Å, $c = 94.88$ Å, and $\alpha = \beta = \gamma = 90°$. A functional dimer was present in the ASU.

Overall structure of ectoine hydroxylases

The core of nonheme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase[8,9,25] is comprised of mainly antiparallel β-strands that fold to form a distorted/squashed barrel-like structure open at one end allowing the substrate to diffuse into the active center. This barrel-like structure is also termed 'cupin fold'.[8,9] The amino acid sequences of the *Vs*EctD and *Sa*EctD proteins described here possess an amino acid sequence identity of 51%. An overall comparison of the five determined EctD crystal structures from *Vs*EctD and *Sa*EctD (PDB accession codes: 3EMR, 4NMI, 4MHR 4MHU, and 4Q5O) revealed a high degree of identity with a root–mean–square deviation (r.m.s.d.) that ranges from 1.3 to 1.6 Å over 279 Cα atoms. An even lower RMSD of 0.5–0.9 Å was found when the three different *Sa*EctD structures or the two *Vs*EctD structures were compared with each other.

As the monomers of all known EctD structures are nearly identical in overall shape, we describe in the following section only the overall fold for the apo-*Sa*EctD protein. The *Sa*EctD crystal structure consists of a double-stranded β-helix core surrounded and stabilized by a number of α-helices (3D Structure). This core, termed the jelly-roll or cupin fold,[8,9] is formed by two four-stranded antiparallel β-sheets that are arranged in the form of a β-sandwich. This type of architecture has previously been observed not only for the EctD proteins[10,11,21,23] but also for many other nonheme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases.[8,9,26] This is highlighted by a DALI search,[27] which revealed structural similarity of EctD to the human phytanoyl-CoA hydroxylases PhyH (PDB code 2A1X; r.m.s.d of 2.6 Å over 252 C-alpha atoms when compared to the *Sa*EctD structure)[28] PhyHD1A (PDB code 2OPW r.m.s.d. of 2.2 Å over 286 Cα atoms),[29] the halogenases SyrB2 from *Pseudomonas syringae* (PDB code 2FCV; r.m.s.d. of 2.9 Å over 299 Cα atoms),[30] CytC3 from a *Streptomyces* soil isolate (PDB code 3GJA; r.m.s.d. of 3.0 Å over 204 C-alpha atoms),[31] and CurA from the cyanobacterium *Lyngbya majuscula* (PDB code 3NNM; r.m.s.d. of 3.3 Å over 222 Cα atoms).[32] While the core of the cupin-type fold of EctD is highly similar to these proteins, the rest of the EctD protein differs from the structurally compared proteins resulting in higher r.m.s.d. values.

Alignments of the amino acid sequences of over 400 ectoine hydroxylases revealed a highly conserved signature sequence motif of 17 amino acids in length (FXWHSDFETWHXEDGM/LP).[1,6,10,11,18] This string of amino acids forms an extended α-helix followed by a short β-strand; it structures one side of the EctD cupin-barrel (highlighted in green in 3D Structure). The signature sequence region is structurally important for the overall fold of the EctD proteins and notably contains five residues that contribute to the binding of all three ligands of the ectoine hydroxylase: Fe(II), 2-oxoglutarate, and ectoine. In the *Sa*EctD protein, these residues comprise the side chains of His144 and Asp146 involved in iron binding, Phe141 involved in 2-oxoglutarate binding, and His144, Asp146, Thr149, and Trp150 that are involved in binding and the correct positioning of the substrate ectoine.

*Sa*EctD contains a large loop (amino acids 191–210), which is disordered and hence not visible in the electron density map. By amino acid sequence alignment, this loop corresponds to residues 195–211 in the *Vs*EctD ectoine hydroxylase and it is disordered in both available crystal structures.[10,11] Interestingly, in the ligand-free crystal

This article was published in the *Encyclopedia of Inorganic and Bioinorganic Chemistry* in 2016 by John Wiley & Sons, Ltd.
DOI: 10.1002/9781119951438.eibc2440

structures of the phytanoyl-CoA hydroxylase PhyH,[28] the asparagine hydroxylase AsnO from *Streptomyces coelicolor*,[33] the L-arginine oxygenase VioC from *Streptomyces vinaceus*,[34] and the taurine dioxygenases from *E. coli* and *Pseudomonas putida*,[35] a similar loop is disordered as well and became only visible in crystal structures with bound substrates. It is thought that this mobile loop functions as a lid that shields the active site form the solvent once the ligands are bound.

Oligomeric state: EctD is a dimeric protein

To determine the oligomeric state of the purified *Sa*EctD in solution, we carried out high-pressure liquid chromatography coupled to multiangle light scattering (HPLC-MALS) analysis. This method allows the calculation of the size of the particle (or protein) running through the column. The normalized elution profiles from the UV, refractive index, and light scattering detectors revealed that the protein solution was homogeneous and monodisperse. After determining protein concentration by the refractive index, a molecular mass of 70.73 ± 1.1 kDa was obtained for the *Sa*EctD protein.[21] This molecular mass corresponds very well with the theoretical molecular mass of a dimer of the recombinant protein (calculated molecular mass of the monomer, including the Strep-tag-II affinity peptide: 35.29 kDa).[21] Hence, *Sa*EctD is a dimer in solution, in line with the result obtained for six other EctD proteins as assessed by conventional size-exclusion chromatography.[11] HPLC-MALS analysis of the *Sa*EctD protein was repeated in the presence of the 2-oxoglutarate and ectoine substrates revealing that *Sa*EctD was a dimer under these conditions as well. Overall, these data suggest that the hydroxylation of ectoine is mediated by a dimeric EctD enzyme.

The crystal structures of *Sa*EctD and *Vs*EctD displayed a different crystal packing (as observed by the different symmetries of the crystals) and protein composition in the ASU of the crystals. Inspection of the crystal packing and analysis of the monomer–monomer interactions in the crystals revealed a physiological relevant dimer (Figure 1).[21] Using the PDBe PISA software,[36] 10 potential amino acids were identified to be involved in monomer–monomer interaction. Each of these is located in loop areas pointing from one monomer toward the other and thereby contributes to the formation of the dimer interface. These amino acids are, however, only moderately conserved, as judged by consulting an alignment of over 400 EctD-type proteins.[6,11]

Iron-binding site in *Sa*EctD and *Vs*EctD

Both *Vs*EctD and *Sa*EctD were crystallized in the Fe(II) bound state. In *Vs*EctD, the iron cofactor is coordinated by the side chains of His146, Asp148, and His248, together with three water molecules that interact with the iron ligand, arranged in an almost perfect octahedral geometry (Figure 3(a)).[10] To probe the functional role of His146, Asp148, and His248 in *Vs*EctD for iron-binding and enzyme activity,[1] Ala mutants were created for each of the three iron-binding residues. These *Vs*EctD variants neither were able to bind iron nor were they enzymatically active. These residues form a structurally strictly conserved HXD/E…H motif, the 2-His-1-carboxylate facial triad, forming a type of mononuclear iron center that can be found in many members of the dioxygenase superfamily.[7–9,26] The structure of *Sa*EctD revealed a set of interactions involved in binding the Fe(II) ion similar to that found in the *Vs*EctD protein (Figure 3(a)). However, in the *Sa*EctD structure, no water molecules were observed near the iron ligand owing to the moderate 2.6 Å resolution of this structure.[21] In other EctD-related protein structures adopting a cupin fold, which were identified by a structural comparison search using the DALI server (Table 2), the Fe(II) ligand is bound *via* a octahedral geometry as well.

Crystal structures of EctD were also obtained in the unliganded apo form. Comparison of the structural arrangement in the apo- and iron-bound structure of *Vs*EctD and *Sa*EctD revealed no significant structural differences. This highlights that the iron ligand does not induce significant structural changes in the ectoine hydroxylase up on binding, rather the iron-binding site is preset of receive the Fe(II) cofactor.[21]

Mechanistic aspects

Unfortunately, no crystal structure of the substrate-bound form of *Vs*EctD has been reported. To get a glimpse of the Fe(II)-binding site in the presence of 2-oxoglutarate, structural alignments were preformed with members of the nonheme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases crystallized in complex with either 2-oxoglutarate or analogs of this cosubstrate.[37,38] This suggested a binding site in *Vs*EctD for this cosubstrate similar to that experimentally determined in other nonheme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases. In this *in silico* model, the 2-oxoglutarate interacts twice with the Fe(II) atom and replaces the water molecules observed in the Fe(II) crystal structure by which the octahedral conformation of the metal remains (Figure 3(b)) and Table 2).

While no 2-oxoglutarate-bound *Vs*EctD crystal structure could be obtained, that of *Sa*EctD was solved with 2-oxoglutarate, and in the presence of the product 5-hydroxyectoine, the so-called dead-end enzyme state. In this structure, the binding site of 2-oxoglutarate was confirmed to be similar to the binding site observed in other 2-oxoglutarate-dependent dioxygenases, including that predicted *in silico* for the *Vs*EctD protein. In addition to the interactions with the side chains of His144, His245,

This article was published in the *Encyclopedia of Inorganic and Bioinorganic Chemistry* in 2016 by John Wiley & Sons, Ltd.
DOI: 10.1002/9781119951438.eibc2440

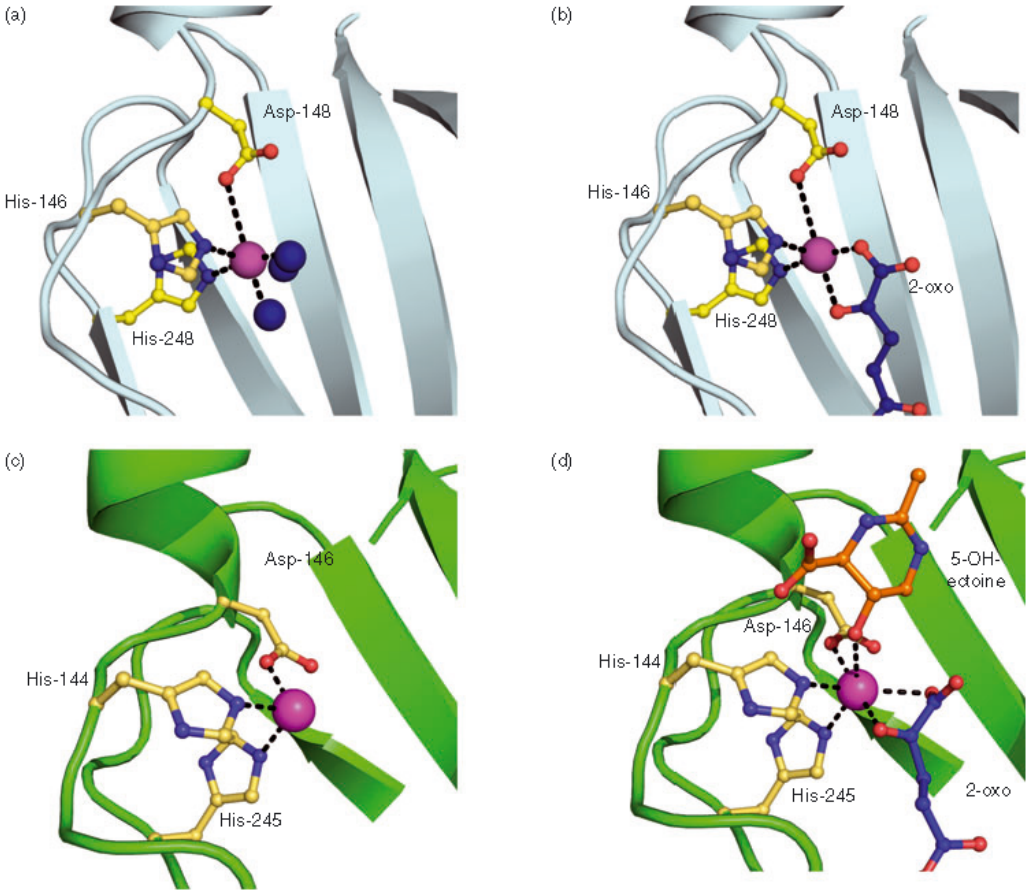


Figure 3 Iron-binding site in *Vs*EctD and *Sa*EctD. (a) The iron-binding site of *Vs*EctD is highlighted. The octahedral coordination of Fe(II) is mediated by His146, Asp148, His248, H_2O-823, H_2O-824, and H_2O-824. All distances of the side chain atoms toward the Fe(II) ligand are in between 2.3 and 2.6 Å. (b) Architecture of the Fe(II)-binding site in the apo-*Sa*EctD structure. The coordination of Fe(II) is mediated by His144, Asp146, and His245. The resolution of this structure does not allow the placing of water molecules and therefore no further interaction partners are observed. All distances of the side chain atoms toward the Fe(II) ligand are in between 2.3 and 2.7 Å. (c) The Fe(II)-binding site in the *Sa*EctD in complex with 2-oxoglutarate and 5-hydroxyectoine is depicted. (d) The octahedral coordination of Fe(II) is mediated by His144, Asp146, His245, 2-oxoglutarate, and 5-hydroxyectoine. All distances of the side chain atoms toward the Fe(II) ligand are in between 2.3 and 2.6 Å. The PDB accession codes used in this figure are 3EMR, 4NMI, 4MHU, and 4Q5O.

and Asp146 and with 2-oxoglutarate, one further interaction appeared to be present in this dead-end complex. Fe(II) interacted with the 5-hydroxy group present in the product 5-hydroxyectoine as expected for an enzyme reaction mechanisms directly involved the Fe(II) ion (Figure 3(b)).

On the basis of the crystal structures, the biochemical studies with the EctD protein, and in analogy with other nonheme-containing iron(II) and 2-oxoglutarate-dependent dioxygenases,[6,7,9,25,39,40] the following enzyme reaction (Figure 4) for the ectoine hydroxylase was proposed: (**I**) The substrate-free EctD enzyme contains Fe(II), sixfold-coordinated by the 2-His-1-carboxylate facial triad formed by His146, Asp148, and His248, and three water molecules.[10] Two of the water molecules are replaced from the iron center upon addition of 2-oxoglutarate. Subsequently, the substrate ectoine is envisioned to bind to the EctD enzyme *via* the molecular interactions gleaned from the *Sa*EctD:5-hydroxyectoine complex. (**II**) Upon substrate binding, dioxygen replaces the third water molecule from the iron center and oxidizes Fe(II) to Fe(IV). (**III**) The formed Fe(IV)-peroxo species attacks with its anion the carbonyl-C atom of 2-oxoglutarate to yield succinate, CO_2, and the ferryl species (Fe(IV)═O).[41] The ferryl species hydroxylates ectoine to (*4S*,*5S*)-5-hydroxyectoine. (**IV**) This

This article was published in the *Encyclopedia of Inorganic and Bioinorganic Chemistry* in 2016 by John Wiley & Sons, Ltd.
DOI: 10.1002/9781119951438.eibc2440

Table 2 Structures of known EctD enzymes and proteins adopting a similar fold as deduces from a DALI search

Enzyme	PDB code	Ligand	Metal-binding site	Binding geometry
*Vs*EctD	3emr	Fe(II)	His146, Asp148, His248, H_2O-823, H_2O-824, H_2O-824	Octahedral
*Sa*EctD	4mhu	Fe(II)	His144, Asp146, His245	Square planar[a]
*Sa*EctD	4q5o	Fe(II), 2-oxo, 5-hydroxyectoine	His144, Asp146, His245, 2oxo, 5hydroxyectoine	Octahedral
PhyH	2a1x	Fe(II), 2-oxo	His175, Asp177, His264, 2oxo, H_2O-463	Octahedral
Avi01	4xaa	Ni(II)	His125, Asp127, His196, H_2O-426, H_2O-429, H_2O-440	Octahedral
Evd01	4xbz	Ni(II)	His139, Asp141, His223, glycerol-402	Octahedral
SyrB2	2fct	Fe(II), 2-oxo	His116, His235, 2-oxo, chloride ion	Octahedral
FtmOx1	4y5s	Fe(II), 2-oxo	His133, Asp135, His209, 2-oxo	Octahedral
HygX	4xca	Ni(II), 2-oxo	His102, His205, 2-oxo, H_2O-458, H_2O-459	Octahedral
CytC3	3gjb	Fe(II), 2-oxo	His118, His240, 2-oxo, H_2O-402, H_2O-400	Octahedral
CurA	3nnf	Fe(III), 2-oxo	His115, His228, cloride-321, 2-oxo, FMT	Octahedral
PtlH	2rdq	Fe(III), 2-oxo	His137, Asp139, His226, 2-oxo, H_2O-270	Octahedral
Phd2	3ouj	Fe(III), 2-oxo	His313, Asp315, His374, 2-oxo, H_2O-23	Octahedral

The binding geometry was calculated using the CMM software.[37]
[a]For the *Sa*EctD structure, a geometry was calculated to be square planar.

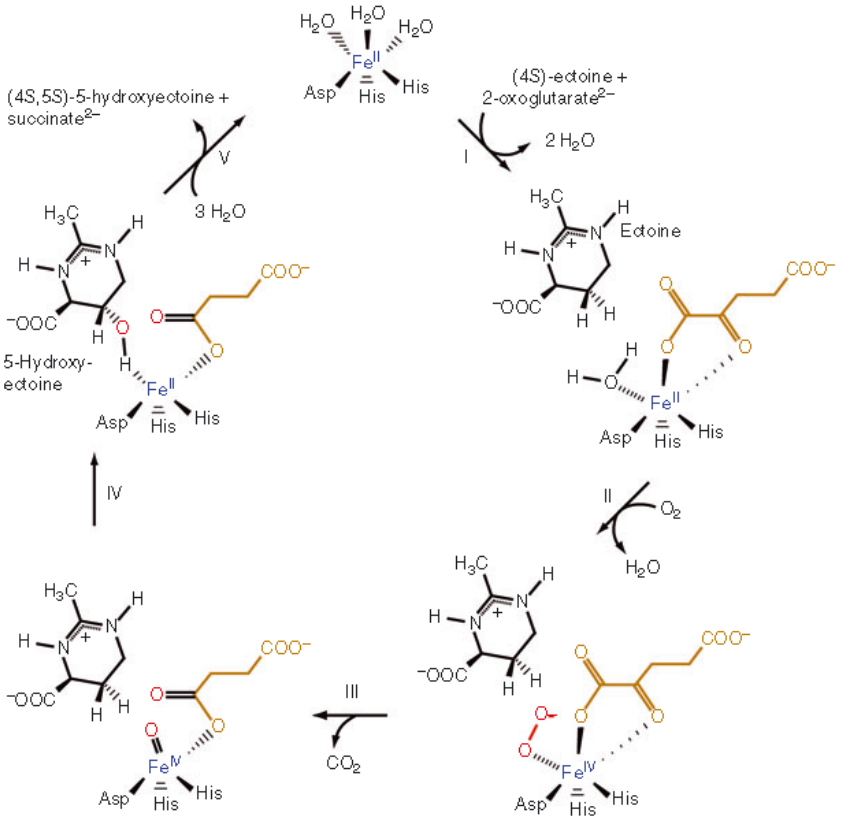


Figure 4 Proposal for the reaction mechanism of the conversion of ectoine into 5-hydroxyectoine by the ectoine hydroxylase.[6] For a detailed explanation see text.

This article was published in the *Encyclopedia of Inorganic and Bioinorganic Chemistry* in 2016 by John Wiley & Sons, Ltd.
DOI: 10.1002/9781119951438.eibc2440

may occur either by abstraction of the pro-*S* hydrogen at C5 by Fe(IV)═O and readdition of the hydroxyl radical from Fe(III)—OH to the ectoine radical intermediate with retention of the configuration (rebound mechanism) or by insertion of the ferryl oxygen atom into the C—H bond (two-state reactivity mechanism). (**V**) Finally, the product (*4S*,*5S*)-5-hydroxyectoine and the coproduct succinate are released. Three water molecules will then occupy again the empty coordination site of Fe(II).

Role of the ectoine hydroxylase

Ectoine and 5-hydroxyectoine are potent osmostress protectants for microorganisms and also serve as protectants against extremes in growth temperature. They can exert these functions both when they are synthesized and when they are imported *via* osmotically controlled transport systems.[13] Increased production of 5-hydroxyectoine often occurs when the growth of osmotically stressed microbial cells slows down and the culture enters stationary phase, indicating that the cytoprotective effects of 5-hydroxyectoine go beyond its well-known function as an osmostress protectant. Indeed, the EctD-catalyzed hydroxylation of ectoine endows the newly formed 5-hydroxyectoine with properties that are often superior to that exhibited by ectoine; for example, 5-hydroxyectoine provides enhanced heat-stress protection[4,18,38] and it is also a very effective protectant against desiccation in comparison with ectoine. This latter property is caused by the ability of 5-hydroxyectoine to form glass-like structures.[38] The question therefore arises, why do not all ectoine-producing bacteria[5,11] also produce 5-hydroxyectoine? Part of the answer can readily be deduced from the biochemical properties of the ectoine hydroxylase as it is a strictly oxygen-dependent enzyme.[1,5,11] Microorganisms with an anaerobic metabolism could therefore not make use of it, as highlighted by the recent study of ectoine-producing members of the *Archaea*.[5] As members of the *Methanosaeta* and *Methanobacterium* genera are all strict anaerobes, the absence of *ectD* in the genomes of these methanogens is readily understandable. On the other hand, the presence of *ectD* genes in *Nitrosopumilus maritimus* SCM1, other members of the *Nitrosopumilus* genus, and related Marine Group I Thaumarchaeota is consistent with the lifestyle of these *Archaea* as oxygen-dependent nitrifying microorganisms.[5]

The crystallographic analysis of the Fe-(*Sa*)EctD/2-oxoglutarate/5-hydroxyectoine complex provided a detailed snapshot of the architecture of the active site of the ectoine hydroxylase. Synthetic ectoine derivatives with either reduced or expanded ring sizes have been reported.[42,43] The architecture of the catalytic core of the EctD protein[21] is probably flexible enough to allow the hydroxylation of synthetic, ectoine-related molecules. An EctD-mediated hydroxylation might endow them with novel stress-protective and structure-preserving functions in the same way that it allows 5-hydroxyectoine to function strikingly different from ectoine in alleviating desiccation stress.[38] As a result, the EctD enzyme is not only of interest from an ecophysiological and biotechnological point of view,[5,11–13] it also has the potential for use in chemical biology.

RELATED ARTICLES

2,3-Dihydroxybiphenyl 1,2-Dioxygenase; Naphthalene 1,2-Dioxygenase; Iron Proteins with Dinuclear Active Sites

REFERENCES

1 J Bursy, AJ Pierik, N Pica and E Bremer, *J Biol Chem*, **282**, 31147–31155 (2007).
2 L Inbar, F Frolow and A Lapidot, *Eur J Biochem*, **214**, 897–906 (1993).
3 J Prabhu, F Schauwecker, N Grammel, U Keller and M Bernhard, *Appl Environ Microbiol*, **70**, 3130–3132 (2004).
4 R Garcia-Estepa, M Argandona, M Reina-Bueno, N Capote, F Iglesias-Guerra, JJ Nieto and C Vargas, *J Bacteriol*, **188**, 3774–3784 (2006).
5 N Widderich, Czech L, F J Elling, M Koenneke, N Stoeveken, M Pittelkow, R Riclea, J S Dickschat, J Heider, E Bremer, *Environ Microbiol*(2015), doi: 10.1111/1462-2920.13156. [Epub ahead of print] PMID: 26636559.
6 N Widderich, M Pittelkow, A Höppner, D Mulnaes, W Buckel, H Gohlke, SH Smits and E Bremer, *J Mol Biol*, **426**(3), 586–600 (2014).
7 RP Hausinger, *Crit Rev Biochem Mol Biol*, **39**, 21–68 (2004).
8 W Aik, MA McDonough, A Thalhammer, R Chowdhury and CJ Schofield, *Curr Opin Struct Biol*, **22**, 691–700 (2012).
9 JA Hangasky, CY Taabazuing, MA Valliere and MJ Knapp, *Metallomics*, **5**, 287–301 (2013).
10 K Reuter, M Pittelkow, J Bursy, A Heine, T Craan and E Bremer, *PLoS One*, **5**, e10647 (2010).
11 N Widderich, A Höppner, M Pittelkow, J Heider, SHJ Smits and E Bremer, *PLoS One*, **9**, e93809 (2014).
12 G Lentzen and T Schwarz, *Appl Microbiol Biotech*, **72**, 623–634 (2006).
13 JM Pastor, M Salvador, M Argandona, V Bernal, M Reina-Bueno, LN Csonka, JL Iborra, C Vargas, JJ Nieto and M Canovas, *Biotechnol Adv*, **28**, 782–801 (2010).
14 K Schwibbert, A Marin-Sanguino, I Bagyan, G Heidrich, G Lentzen, H Seitz, M Rampp, SC Schuster, HP Klenk, F Pfeiffer, D Oesterhelt and HJ Kunte, *Environ Microbiol*, **13**, 1973–1994 (2011).
15 P Louis and EA Galinski, *Microbiology*, **143**, 1141–1149 (1997).
16 AU Kuhlmann and E Bremer, *Appl Environ Microbiol*, **68**, 772–783 (2002).
17 AU Kuhlmann, J Bursy, S Gimpel, T Hoffmann and E Bremer, *Appl Environ Microbiol*, **74**, 4560–4563 (2008).

This article was published in the *Encyclopedia of Inorganic and Bioinorganic Chemistry* in 2016 by John Wiley & Sons, Ltd.
DOI: 10.1002/9781119951438.eibc2440

The ectoine hydroxylase: a nonheme-containing iron(II) and 2-oxoglutarate-dependent dioxygenase

18 J Bursy, AU Kuhlmann, M Pittelkow, H Hartmann, M Jebbar, AJ Pierik and E Bremer, *Appl Environ Microbiol*, **74**, 7286–7296 (2008).

19 SH Saum and V Muller, *Environ Microbiol*, **10**, 716–726 (2008).

20 II Mustakhimov, AS Reshetnikov, AS Glukhov, VN Khmelenina, MG Kalyuzhnaya and YA Trotsenko, *J Bacteriol*, **192**, 410–417 (2010).

21 A Hoppner, N Widderich, M Lenders, E Bremer and SH Smits, *J Biol Chem*, **289**, 29570–29583 (2014).

22 W Lovenberg, BB Buchanan and JC Rabinowitz, *J Biol Chem*, **238**, 3899–3913 (1963).

23 A Hoeppner, N Widderich, E Bremer and SH Smits, *Acta Crystallogr F Struct Biol Commun*, **70**, 493–496 (2014).

24 A Hoeppner, L Schmitt, S H Smits, Proteins and Their Ligands: Their Importance and How to Crystallize Them. Intech Books, Vol. **II**, pp. 44 (2013), Available from: http://www.intechopen.com/books/advanced-topics-on-crystal-growth/proteins-and-their-ligands-their-importance-and-how-to-crystallize-them.

25 AG Prescott and MD Lloyd, *Nat Prod Rep*, **17**, 367–383 (2000).

26 S Kundu, *BMC Res Notes*, **5**, 410 (2012).

27 L Holm and P Rosenström, *Nucl Acids Res*, **38**, W545–W549 (2010).

28 MA McDonough, KL Kavanagh, D Butler, T Searls, U Oppermann and CJ Schofield, *J Biol Chem*, **280**, 41101–41110 (2005).

29 Z Zhang, GT Kochan, SS Ng, KL Kavanagh, U Oppermann, CJ Schofield and MA McDonough, *Biochem Biophys Res Commun*, **408**, 553–558 (2011).

30 LC Blasiak, FH Vaillancourt, CT Walsh and CL Drennan, *Nature*, **440**, 368–371 (2006).

31 C Wong, DG Fujimori, CT Walsh and CL Drennan, *J Am Chem Soc*, **131**, 4872–4879 (2009).

32 D Khare, B Wang, L Gu, J Razelun, DH Sherman, WH Gerwick, K Hakansson and JL Smith, *Proc Natl Acad Sci U S A*, **107**, 14099–14104 (2010).

33 M Strieker, F Kopp, C Mahlert, LO Essen and MA Marahiel, *ACS Chem Biol*, **2**, 187–196 (2007).

34 V Helmetag, SA Samel, MG Thomas, MA Marahiel and LO Essen, *FEBS J*, **276**, 3669–3682 (2009).

35 SH Knauer, O Hartl-Spiegelhauer, S Schwarzinger, P Hanzelmann and H Dobbek, *FEBS J*, **279**, 816–831 (2012).

36 E Krissinel and K Henrick, *J Mol Biol*, **372**, 774–797 (2007).

37 H Zheng, MD Chordia, DR Cooper, M Chruszcz, P Muller, GM Sheldrick and W Minor, *Nat Protoc*, **9**, 156–170 (2014).

38 C Tanne, EA Golovina, FA Hoekstra, A Meffert and EA Galinski, *Front Microbiol*, **5**, 150 (2014).

39 GD Straganz and B Nidetzky, *Chembiochem*, **7**, 1536–1548 (2006).

40 PK Grzyska, EH Appelman, RP Hausinger and DA Proshlyakov, *Proc Natl Acad Sci U S A*, **107**, 3982–3987 (2010).

41 PJ Riggs-Gelasco, JC Price, RB Guyer, JH Brehm, EW Barr, JM Bollinger Jr, and C Krebs, *J Am Chem Soc*, **126**, 8108–8109 (2004).

42 M Schnoor, P Voss, P Cullen, T Boking, HJ Galla, EA Galinski and S Lorkowski, *Biochem Biophys Res Commun*, **322**, 867–872 (2004).

43 EM Witt, NW Davies and EA Galinski, *Appl Microbiol Biotechnol*, **91**, 113–122 (2011).

This article was published in the *Encyclopedia of Inorganic and Bioinorganic Chemistry* in 2016 by John Wiley & Sons, Ltd.
DOI: 10.1002/9781119951438.eibc2440

II. Publication 7

The seventh manuscript is entitled 'Salt-sensitivity of σ^H and Spo0A prevents sporulation of *Bacillus subtilis* at high osmolarity avoiding death during cellular differentiation'. It is published in 'Molecular Microbiology' under the following doi: 10.1111/mmi.13304.

Widderich, N., Rodrigues, C.D.A., Commichau, F.M., Fischer, K.E., Ramirez-Guadiana, F.H., Rudner, D.Z. and Bremer, E. Salt-sensitivity of σ^H and Spo0A prevents sporulation of *Bacillus subtilis* at high osmolarity avoiding death during cellular differentiation. *Mol Microbiol*. (*in press*). doi: 10.1111/mmi.13304.

This paper addresses the question why sporulation is blocked in starving *Bacillus subtilis* cells that simultaneously face osmotic stress. It is demonstrated that spore formation is drastically impaired in starving high-salinity challenged *B. subtilis* cells and that the block occurs prior to asymmetric division. Further, it reveals that transcription dependent on σ^H and Spo0A is impaired and that association of σ^H with core-RNA-polymerase is reduced at high salinity. Additionally, it is demonstrated that cells induced to bypass the early block are impaired in later steps of the sporulation pathway and frequently undergo lysis prior to the formation of a mature spore indicating that the block in sporulation serves as a safeguard to avert osmotically stressed *B. subtilis* cells from committing to a developmental program they cannot complete and would die trying.

Personal contribution:
I conducted physiological experiments and performed immunoblot blot analysis and fluorescence microscopy experiments (together with Christopher Rodrigues). I further conducted pull-down experiments and enriched for suppressor strains. I analyzed the data (together with Christopher Rodrigues, David Rudner and Erhard Bremer) and prepared Figures and Table (together with Christopher Rodrigues). I contributed to the writing of the manuscript together with Christopher Rodrigues, David Rudner and my PhD supervisor Erhard Bremer.

Molecular Microbiology (2016) 00(00), 00–00 ■
doi: 10.1111/mmi.13304
First published online 00 Month 2016

Salt-sensitivity of σ^H and Spo0A prevents sporulation of *Bacillus subtilis* at high osmolarity avoiding death during cellular differentiation

Nils Widderich,[1†] Christopher D. A. Rodrigues,[2†] Fabian M. Commichau,[3] Kathleen E. Fischer,[1] Fernando H. Ramirez-Guadiana,[2] David Z. Rudner[2*] and Erhard Bremer[1*]

[1]*Department of Biology, Laboratory for Molecular Microbiology, Philipps-University Marburg, Karl-von-Frisch Str. 8, D-35043 Marburg, Germany.*

[2]*Department of Microbiology and Immunobiology, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115-5701, USA.*

[3]*Department of General Microbiology, Institute of Microbiology and Genetics, Georg August University Göttingen, Griesebachstr, 8, D-37077 Göttingen, Germany.*

Summary

The spore-forming bacterium *Bacillus subtilis* frequently experiences high osmolarity as a result of desiccation in the soil. The formation of a highly desiccation-resistant endospore might serve as a logical osmostress escape route when vegetative growth is no longer possible. However, sporulation efficiency drastically decreases concomitant with an increase in the external salinity. Fluorescence microscopy of sporulation-specific promoter fusions to *gfp* revealed that high salinity blocks entry into the sporulation pathway at a very early stage. Specifically, we show that both Spo0A- and SigH-dependent transcription are impaired. Furthermore, we demonstrate that the association of SigH with core RNA polymerase is reduced under these conditions. Suppressors that modestly increase sporulation efficiency at high salinity map to the coding region of *sigH* and in the regulatory region of *kinA*, encoding one the sensor kinases that activates Spo0A. These findings led us to discover that *B. subtilis* cells that overproduce KinA can bypass the salt-imposed block in sporulation. Importantly, these cells are impaired in the morphological process of engulfment and late forespore gene expression and frequently undergo lysis. Altogether our data indicate that *B. subtilis* blocks entry into sporulation in high-salinity environments preventing commitment to a developmental program that it cannot complete.

Accepted 4 December, 2015. *For correspondence. E-mail bremer@staff.uni-marburg.de; Tel. (+49)-6421-2821529; Fax (+49)-6421-2828979 and E-mail rudner@hms.harvard.edu; Tel. (+1)-617-432-4455; Fax (+1)-617-432-4787. †These authors contributed equally to this work.

Introduction

The soil is a challenging habitat for microorganisms. Both, the supply of nutrients and physical parameters like pH, temperature and osmolarity fluctuate frequently. *Bacillus subtilis* is a soil-dwelling bacterium (Mandic-Mulec *et al.*, 2015) that is well adapted to life in the upper layers of the soil (Earl *et al.*, 2008; Belda *et al.*, 2013). It counteracts various challenges through the induction of the σ^B-controlled general stress regulon (Hecker *et al.*, 2007; Price, 2011) and the engagement of stress-specific adaptation mechanisms (Sonenshein *et al.*, 2002). Furthermore, in response to nutrient limitation, *B. subtilis* can differentiate into a dormant and stress-resistant endospore (Errington, 2003; Higgins and Dworkin, 2012; Setlow, 2014). Here, we have investigated how *B. subtilis* simultaneously responds to two common challenges that it encounters in its native habitat: starvation and high salinity (Bremer and Krämer, 2000; Sonenshein, 2000; Bremer, 2002; Higgins and Dworkin, 2012).

In response to starvation, *B. subtilis* initiates a highly orchestrated developmental program involving hundreds of genes (Errington, 2003; Eichenberger *et al.*, 2004; Steil *et al.*, 2005; Higgins and Dworkin, 2012; Nicolas *et al.*, 2012). During this process, the cell undergoes an asymmetric division resulting in two cell types, the mother cell and forespore, which follow distinct developmental programs of gene expression driven by stage- and compartment-specific transcription factors. These programs are linked to each other through cell–cell signaling pathways, ensuring that

gene expression in one cell is kept in register with gene expression in the other. Upon maturation of the developing forespore, lysis of the mother cell releases a highly stress-resistant endospore into the environment (Stragier and Losick, 1996; Errington, 2003; Higgins and Dworkin, 2012). *B. subtilis* spores can remain dormant in a given ecosystem for extended periods of time, yet can rapidly germinate in response to specific germinants that signal conditions are conducive for vegetative growth (Setlow, 2014; Sinai *et al.*, 2015; Sturm and Dworkin, 2015).

Two transcription factors govern the entry into the sporulation pathway: the response regulator Spo0A and the alternative sigma factor σ^H. These regulators control the expression of hundreds of genes involved in stationary phase adaptation in addition to those involved in the earliest stages of spore formation. Both transcription factors are under complex and interconnected regulatory control (Hoch, 1993; Grossman, 1995; Sonenshein, 2000; Phillips and Strauch, 2002; Chastanet *et al.*, 2010). A set of sensor kinases (KinA – KinE) trigger activation of Spo0A, the master regulator of sporulation, via a multicomponent phosphorelay (Burbulys *et al.*, 1991; Stragier and Losick, 1996). In most cases, the specific signals sensed by these kinases are incompletely understood. Phosphorylated Spo0A (Spo0A~P) directly and indirectly regulates the transcription of over 400 genes (Molle *et al.*, 2003). Depending on the flux through the phosphorelay, different levels of Spo0A~P are generated allowing a graded adaptive response since Spo0A-controled promoters have different affinities for Spo0A~P (Chung *et al.*, 1994; Fujita *et al.*, 2005). The sporulation genes under Spo0A control have low-affinity Spo0A binding sites and therefore require high concentrations of Spo0A~P for their effective expression (Fujita *et al.*, 2005). One of the genes regulated indirectly by Spo0A~P is *sigH* (also referred to as *spo0H*), which encodes the alternative sigma factor σ^H (Haldenwang, 1995; Stragier and Losick, 1996; Errington, 2003; Higgins and Dworkin, 2012). The globally acting repressor AbrB negatively regulates transcription of *sigH*, and Spo0A, in turn, represses *abrB* transcription. It does so at low Spo0A~P concentrations thereby resulting in an early increase in σ^H levels (Strauch *et al.*, 1990; Fujita *et al.*, 2005). The elevated cellular level of σ^H then promotes expression of sporulation genes, including *spo0A*, *spo0F* (encoding one of the phosphorelay components) and *kinA*, the gene for the central sporulation-specific sensor kinase KinA. In addition, Spo0A~P activates transcription of *spo0A* and *spo0F* genes (Britton *et al.*, 2002; Chastanet and Losick, 2011). These positive feedback loops, in combination with finely tuned flux through the phosphorelay, increase Spo0A and σ^H levels and activity, resulting in the expression of early sporulation genes and entry into the sporulation pathway (Predich *et al.*, 1992; Britton *et al.*, 2002; Chastanet and Losick, 2011).

In its soil habitat (Earl *et al.*, 2008; Mandic-Mulec *et al.*, 2015), *B. subtilis* not only encounters starvation (Sonenshein, 2000; Errington, 2003), but desiccation also subjects it to high salinity (osmolarity) (Bremer and Krämer, 2000; Bremer, 2002). The cell responds to such osmotic challenges through increased potassium import (Whatmore *et al.*, 1990; Holtmann *et al.*, 2003), the synthesis of the compatible solutes proline and glycine betaine (Whatmore *et al.*, 1990; Brill *et al.*, 2011; Nau-Wagner *et al.*, 2012), and the uptake of various osmoprotectants (Hoffmann *et al.*, 2013; Zaprasis *et al.*, 2013; Bashir *et al.*, 2014; Broy *et al.*, 2015). These cellular adjustments prevent efflux of water, maintain turgor at physiologically adequate levels and optimize the solvent properties of the cytoplasm for vital biochemical reactions (Bremer and Krämer, 2000; Bremer, 2002; Wood, 2011). As a result, growth of *B. subtilis* can continue under otherwise osmotically unfavourable circumstances (Boch *et al.*, 1994).

The pronounced desiccation-resistant properties of the *B. subtilis* endospores (Setlow, 2014) suggest that spore formation would be an effective escape route from adverse high-salinity conditions (Bremer, 2002). However, contrary to expectations, high salinity efficiently blocks sporulation (Kunst and Rapoport, 1995; Ruzal and Sanchez-Rivas, 1998; Ruzal *et al.*, 1998). Work from Sanchez-Rivas and coworkers indicates that this block is imposed relatively early in the sporulation process (Ruzal *et al.*, 1998). However, the specific stage at which this occurs and the underlying molecular mechanism(s) have remained elusive.

Here, we have addressed these ecologically relevant questions and show that *B. subtilis* cells that are continuously exposed to high-salt environments are blocked at the earliest possible stage of sporulation. They fail to activate genes under the control of σ^H and Spo0A~P. Synthetic over-production of KinA resulted in a nearly complete bypass of the block in sporulation initiation in salt-stressed cells. However, these cells were impaired in engulfment, late forespore gene expression and exhibited a high frequency of lysis. Collectively, our data suggest that the molecular block averting entry into sporulation under salt stress is not an accident. Rather, it appears to serve a physiologically important role by preventing osmotically stressed cells from committing to a developmental program they cannot complete. This trait is conserved among domesticated and non-domesticated strains of *B. subtilis*.

Results

High salinity drastically impairs spore formation

During the course of our studies on how *B. subtilis* copes with osmotic stress (Bremer and Krämer, 2000; Bremer, 2002), we found that cells grown for more than

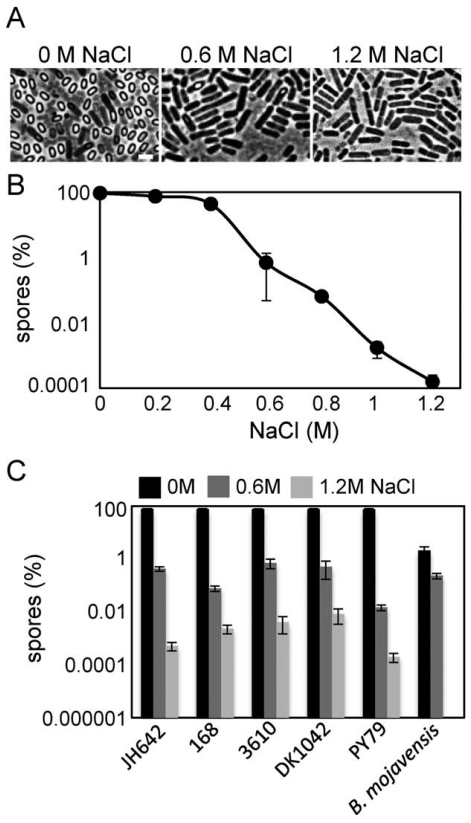


Fig. 1. Sporulation is inhibited by high salinity.
A. Representative phase-contrast images of cells of wild-type *B. subtilis* cells (JH642) sporulated by nutrient exhaustion for 36 h in the presence of the indicated concentrations of NaCl. The scale bar indicates 2 μm.
B. Quantitative assessment of increasing salt concentration on the formation of heat-resistant spores. Sporulation efficiency (spores %) indicates that number of heat-resistant CFU in the indicated concentration of NaCl compared to heat-resistant CFU in the absence of NaCl.
C. Bar graph showing the negative effects of 0.6 M and 1.2 M NaCl on sporulation of the indicated domesticated and nondomesticated *B. subtilis* strains and of *B. mojavensis*.

36 h in Difco Sporulation Medium (DSM) containing 1.2 M NaCl failed to produce phase bright spores (Fig. 1A). This was in stark contrast to the same culture conditions in medium lacking salt (0 M NaCl), in which almost all of the cells formed spores (Fig. 1A). To quantitatively assess the negative effect of high salinity on sporulation, we grew cells of the *B. subtilis* wild-type strain JH642 in DSM in the presence of various concentrations of NaCl and determined the frequency of heat-resistant spores using a standard heat-kill assay. In this assay, the starved cultures are incubated at 80°C for 20 min to kill vegetative and sporulation deficient cells and the number of heat-resistant colony-forming units (CFUs) that result from the outgrowth of the surviving spores are compared to total CFUs (and separately to heat-resistant CFUs in the absence of salt). Increasing salinity reduces growth rate (Boch *et al.*, 1994) (Supporting Information Fig. S1) but the impact was less than twofold at the highest concentration of NaCl (1.2 M) used here. Cells grown in DSM in the absence of salt had a sporulation efficiency of ~90%. A systematic increase in the salt concentration (0.2 M, 0.4 M, 0.6 M NaCl) decreased this value to approximately 70%, 40% and 1%, respectively (Fig. 1B). An additional increase in the salinity led to an even further reduction in sporulation efficiency. Only about one in a million cells grown in the presence of 1.2 M NaCl formed a heat-resistant spore (Fig. 1B). For the remainder of this study, we focused our analysis on media containing either 0.6 M or 1.2 M NaCl.

To investigate whether the inhibition of sporulation by high salinity was specific to the wild-type strain JH642 (Smith *et al.*, 2014), we examined the sporulation efficiency of both domesticated and undomesticated *B. subtilis* isolates. Although sporulation efficiency in the presence of salt was somewhat variable, all strains tested displayed salt sensitivity (Fig. 1C). We also investigated the sporulation efficiency of *Bacillus mojavensis*, a species that is indigenous to the Mojave Desert (Earl *et al.*, 2012). We anticipated that this *Bacillus* species would have an increased tolerance to high salinity as a result of its adaptation to the drought cycles that it experiences in its desert habitat. However, sporulation of *B. mojavensis* was even more sensitive to high salinity than *B. subtilis*, as the presence of 1.2 M NaCl in the DS medium abolished sporulation entirely (Fig. 1C). Altogether, we conclude that the sporulation process is highly sensitive to even modest increases in the salinity of the environment and that this is a common trait of both domesticated and undomesticated *B. subtilis* strains.

High salinity blocks sporulation at a very early stage

Previous studies using electron microscopy to assess the stage at which sporulation is blocked in high salinity suggested that differentiation is inhibited after polar division (Ruzal *et al.*, 1998). To examine a larger number of cells and multiple time points, we used fluorescence microscopy and visualized the morphological stages of sporulation with the membrane dye TMA-DPH. In the same experiment, we monitored the activity of the first forespore-specific transcription factor σ^F (Higgins and

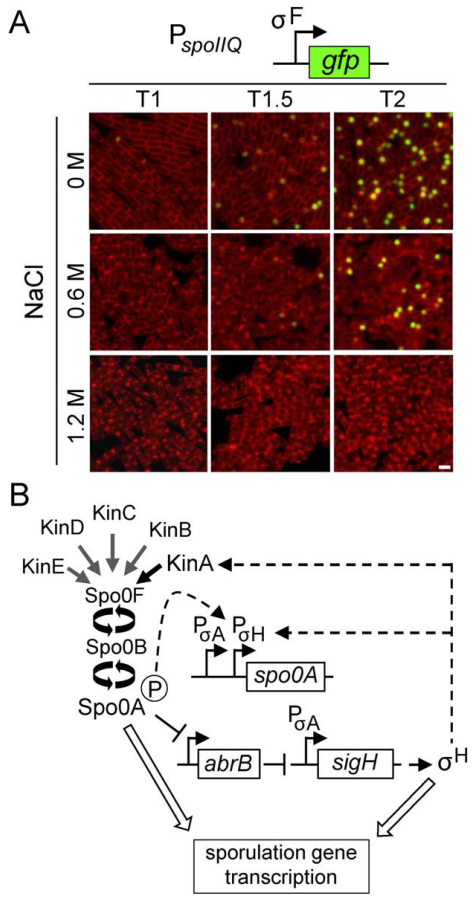


Fig. 2. High salinity inhibits sporulation prior to asymmetric cell division.
A. Representative images of cells (strain BDR1048; Supporting Information Table S1) harbouring a transcriptional reporter (P_{spoIIQ}-*gfp*) for the first forespore-specific transcription factor σ^F induced to sporulate in the indicated concentrations of NaCl. Membranes (false-coloured in red) were stained with TMA-DPH and σ^F-dependent expression is shown in green. Hour 1 (T1), 1.5 (T1.5) and 2 (T2) are shown. Additional time points can be found in Supporting Information Fig. S2. Scale bar is 2 μm.
B. Diagram illustrating part of the complex genetic circuitry involved in the activation of the transcriptional regulators σ^H and Spo0A for sporulation initiation.

Dworkin, 2012) using a P_{spoIIQ}-*gfp* promoter fusion (Londono-Vallejo *et al.*, 1997). We imaged the cells during a sporulation time-course at 30 min intervals after the onset of sporulation (T0) in cultures grown in the presence or absence of NaCl (Fig. 2A and Supporting Information Fig. S2). As reported previously, in the absence of salt >60% of the cells had a polar septum by hour 2 (T2) and most had initiated the process of engulfment (Rodrigues *et al.*, 2013). Furthermore, virtually all the sporulating cells exhibited σ^F-dependent GFP fluorescence in the forespore compartment (Fig. 2A). In the presence of 0.6 M NaCl, 21% ($n > 700$) of the cells had a polar septum and displayed expression of the σ^F reporter. However, in media with 1.2 M NaCl, virtually no cells exhibited asymmetric septation or GFP fluorescence even after 3 h of sporulation (Fig. 2A and Supporting Information Fig. S2). Thus, under our assay conditions, high salinity inhibits sporulation prior to polar division and before the activation of the first compartment-specific transcription factor, σ^F.

Spo0A and σ^H activity are impaired in the presence of high salt

Based on our cytological analysis, we sought to investigate the effects of high salinity on steps in the sporulation process that occur prior to the activation of σ^F. Accordingly, we examined the activity of the transcription factors Spo0A and σ^H that together govern entry into sporulation (Fig. 2B). First, we monitored the expression of a Spo0A-responsive reporter (P_{spoIIE}-*gfp*). The P_{spoIIE} promoter is recognized by σ^A but also requires active Spo0A (Spo0A~P) for its expression (York *et al.*, 1992; Molle *et al.*, 2003). In the absence of salt, Spo0A activity could be detected with this reporter during the first hour (T1) of sporulation and it continued to rise during asymmetric cell division and engulfment (Fig. 3A and Supporting Information Fig. S3). In media containing 0.6 M NaCl, a subset of cells displayed Spo0A-dependent transcription that was similar to the activity observed in the absence of salt. Furthermore, these 'Spo0A high' cells frequently went on to form polar septa (Fig. 3A). However, the majority of cells grown in the presence of 0.6 M NaCl exhibited low Spo0A-dependent transcription (Fig. 3A and Supporting Information Fig. S3). In the presence of 1.2 M NaCl, all cells in the field had virtually undetectable Spo0A activity (Fig. 3A and Supporting Information Fig. S3) consistent with the absence of polar septa (Supporting Information Figs S2B and S4). Thus, salt-stressed *B. subtilis* cells are unable to accumulate sufficient levels of phosphorylated Spo0A to promote the initiation of sporulation. Importantly, high salt did not impair GFP fluorescence as a σ^A-responsive promoter (P_{veg}) fused to *gfp* had strong fluorescence in the presence and absence of NaCl during sporulation (Supporting Information Fig. S5). This result also indicates that high salinity does not globally impair transcription in sporulation medium.

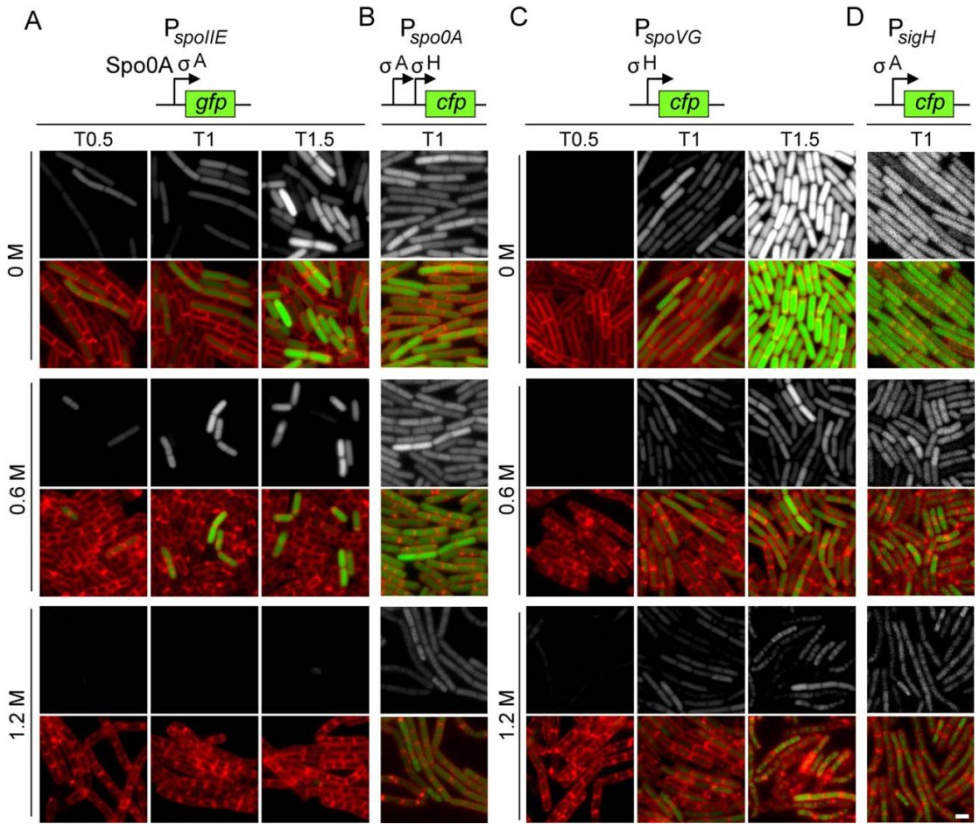


Fig. 3. Spo0A and σ^H activity are impaired in *B. subtilis* cells induced to sporulate at high salinity.
A. Representative images of cells (strain BDR2128; Supporting Information Table S1) harbouring a transcriptional reporter (P_{spoIIE}-*gfp*) monitoring Spo0A-responsive activity that were induced to sporulate in the indicated concentrations of NaCl. For each condition and time point (in hours), GFP fluorescence is shown in black/white and a merged image of membranes (red) and GFP (green) is shown below. The relevant transcription factors that act on the P_{spoIIE} promoter are indicated.
B. Representative images of cells (strain BDR3080; Supporting Information Table S1) harbouring a fusion of the *spo0A* promoter to *cfp* (P_{spo0A}-*cfp*) in the indicated concentrations of NaCl at hour 1 (T1) of sporulation.
C. Representative images of cells (strain BDR3064; Supporting Information Table S1) harbouring a transcriptional reporter (P_{spoVG}-*cfp*) for σ^H-dependent activity that were induced to sporulate in the indicated concentrations of NaCl. (D) Representative images of cells (strain BDR3090; Supporting Information Table S1) harbouring a fusion of the *sigH* promoter to *cfp* (P_{sigH}-*cfp*) at hour 1 of sporulation in the presence of the indicated concentrations of NaCl. For each reporter, all images were scaled identically. Scale bar indicates 2 μm.

To investigate whether transcription of the *spo0A* gene itself was reduced in high salinity, we examined the activity of the *spo0A* promoter using a P_{spo0A}-*cfp* reporter (Fig. 3B). *spo0A* transcription was modestly reduced in the presence 0.6 M NaCl and more severely at 1.2 M. The *spo0A* gene contains two promoters that are inversely controlled by Spo0A. The first promoter (P_v) is recognized by σ^A and negatively regulated by Spo0A~P while the second promoter (P_s) is recognized by σ^H and activated by Spo0A~P (Predich *et al.*, 1992; Strauch *et al.*, 1992; Eymann *et al.*, 2001; Chastanet and Losick, 2011). Thus, the reduction in *spo0A* expression could be due to the salt sensitivity of Spo0A or σ^H.

To investigate whether σ^H activity was influenced by high salt, we analyzed a σ^H-dependent promoter fusion (P_{spoVG}-*cfp*) (Segall and Losick, 1977; Zuber and Losick, 1987). The P_{spoVG} promoter is negatively regulated by AbrB, which is itself repressed by Spo0A at the transcriptional level. However, the extent of AbrB-mediated repression is relatively modest compared to the absolute

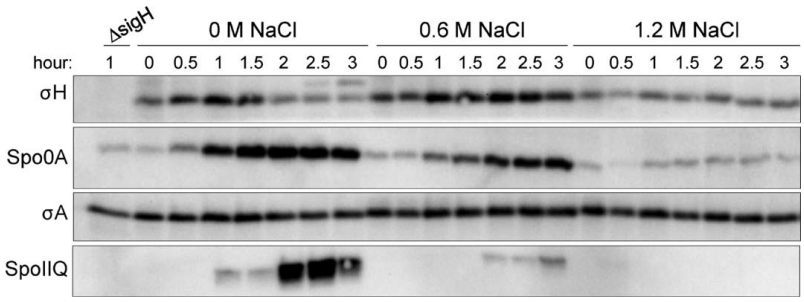


Fig. 4. Spo0A and σ^H protein levels are affected by high salinity. Immunoblot analysis assessing the levels of Spo0A, σ^H and σ^A in cells of the wild-type strain PY79 that were induced to sporulate in the presence of the indicated concentrations of NaCl. Forespore-specific σ^F activity was assessed by the accumulation of SpoIIQ. A *sigH* null mutant (strain BDR3057; Supporting Information Table S1) was included to show the requirement for σ^H activity on the accumulation of Spo0A.

requirement for σ^H (Zuber and Losick, 1987; Liu and Zuber, 2000) (Supporting Information Fig. S6). Accordingly, P_{spoVG}-*cfp* serves as a reliable reporter for σ^H activity. In cells grown in the absence of salt, σ^H-dependent gene expression was detectable at hour 1 (T1) of sporulation and increased significantly over the next half hour leading up to the formation of asymmetric septa (Fig. 3C). In the presence of 0.6 M and 1.2 M NaCl, the level of σ^H activity was slightly reduced at hour 0.5 and failed to increase over the next half hour (Fig. 3C). At hour 1.5, σ^H activity in cells grown in the presence of NaCl was significantly lower than the pre- and post-divisional cells at hour 1.5 that were sporulated in the absence of salt. Similar results were obtained using the P_{spoIIA} promoter fused to GFP (Supporting Information Fig. S7). However, since this promoter is under the control of both σ^H and Spo0A, the impact of high salinity was even more pronounced. Finally, transcription of the gene (*sigH*) encoding the σ^H sigma factor, which is indirectly controlled by Spo0A, was modestly reduced in the presence 0.6 M NaCl and more severely at 1.2 M (Fig. 3D). Collectively, these data suggest that σ^H and Spo0A activity and/or levels are reduced in the presence of high salinity resulting in a failure of *B. subtilis* cells to efficiently enter the sporulation pathway.

Spo0A and σ^H levels are differentially affected by high salinity

To directly assess the levels Spo0A, σ^H and σ^A we examined all three proteins by immunoblot analysis in a sporulation time course. Wild-type cells were induced to sporulate in the presence and absence of NaCl and samples were taken every half hour (Fig. 4). The levels of σ^A remained unchanged in all conditions throughout the entire sporulation time course. As previously reported (Fujita and Sadaie, 1998b,c; Liu and Zuber, 2000), in the absence of NaCl, σ^H levels began to increase within 30 min after the initiation of sporulation and achieved maximum levels 30–60 min later (Fig. 4). By hour 2, σ^H had returned to pre-sporulation levels. In the absence of salt, Spo0A levels closely followed the increase in σ^H, consistent with the requirement of σ^H for *spo0A* expression during sporulation (Predich *et al.*, 1992; Strauch *et al.*, 1992). However, unlike σ^H, Spo0A protein levels remained high for the 3 h time course. This is in line with the observation that Spo0A remains active in the mother cell at later times during the sporulation process (Fujita and Losick, 2003; 2005) (Supporting Information Fig. S3). In the presence of 0.6 M NaCl, σ^H levels increased in a manner similar to the no-salt condition (Fig. 4). This result and our analysis of σ^H activity using P_{spoVG}-*cfp* (Fig. 3C) suggest that σ^H activity is sensitive to high salinity. Interestingly, σ^H levels remained high for the rest of the time course in the presence of 0.6 M NaCl. The nature of this stabilization is currently unknown. However, comparing σ^H activity (assessed by P_{spoVG}-*cfp* fluorescence) and σ^H protein level in the presence and absence of 0.6 M NaCl at hour 2 reinforces the conclusion that σ^H activity is impaired in high salinity (Supporting Information Fig. S8). In contrast to σ^H, Spo0A protein levels increased more slowly in cells grown in the presence of 0.6 M NaCl, and never achieved as high levels as the no salt growth condition (Fig. 4). This is consistent with the salt-sensitivity of σ^H and the reduced activity Spo0A (Fig. 3A and Supporting Information Fig. S3). A comparison of *spo0A* transcription (Fig. 3B) and Spo0A accumulation (Fig. 4) suggests that Spo0A stability is also influenced by high salinity. Finally, both σ^H and Spo0A levels were

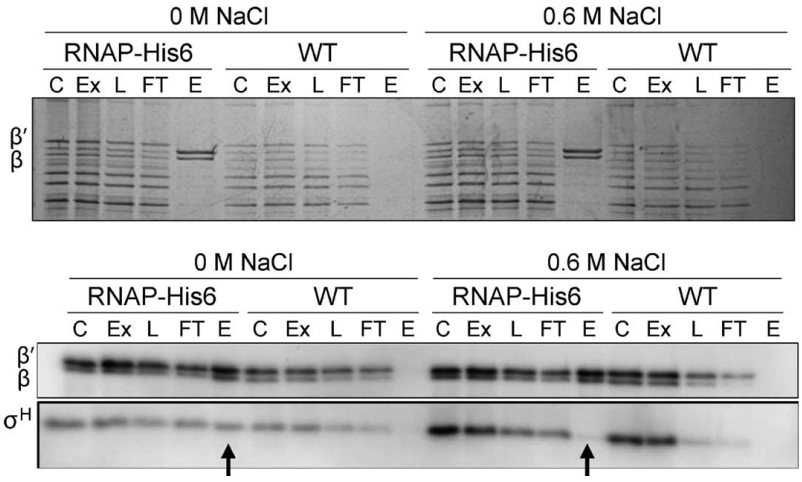


Fig. 5. The association of σ^H with core-RNA-polymerase is affected by high-salinity. Copurification of σ^H with core RNA polymerase from cells induced to sporulate in the indicated concentrations of NaCl. Top panel: Ni^{2+}-chelate affinity purification of β and β'-his6 (RNAP-His6) subunits from sporulating cells at hour 1.5. Whole-cell lysate [C], crude extract [Ex], load [L], flow-through [FT] and eluate [E] from the wild-type (WT) strain PY79 and from strain BDR485 (Supporting Information Table S1; RNAP-His6) were separated by SDS-polyacrylamide gel electrophoresis and stained with Coomassie brilliant blue. Bottom panel: Immunoblot analysis of the same fractions used in the top panel using anti-core and anti-σ^H antibodies. Arrows highlight the relevant elution fractions for comparison.

significantly reduced and failed to accumulate in medium containing 1.2 M NaCl. Collectively, our data indicate that σ^H activity and Spo0A accumulation and activity are sensitive to high salinity and prevent cells from committing to the sporulation pathway.

The interaction between σ^H and core RNA polymerase is impaired in high salinity

We investigated whether the salt-sensitivity of σ^H *in vivo* was related to its inability to stably associate with core RNA-polymerase. To do this, we took advantage of a strain harbouring a functional hexa-histidine fusion to the beta prime subunit of RNA-polymerase (Fujita and Sadaie, 1998b,c). Cells were sporulated by resuspension in the presence and absence of 0.6 M NaCl. At hour 1.5 the cells were harvested by centrifugation and lysed in the presence of 300 mM NaCl followed by Ni^{2+}-NTA affinity purification of RNA-polymerase (see Materials and methods). The amount of copurified σ^H was then determined by immunoblot analysis (Fig. 5). In support of the idea that σ^H is impaired in its ability to associate with core RNA-polymerase in cells grown at high salinity, the amount of copurified σ^H was reduced in cells sporulated in medium containing 0.6 M NaCl compared to cells sporulated in the absence of salt. Importantly, the amount of core RNA-polymerase purified from the two extracts was similar and the level of σ^H in the lysate from the salt-sporulated cells was higher than from the no-salt cells (Fig. 5). We conclude from this set of experiments that the association of σ^H with core RNA polymerase is impaired during sporulation under high salinity conditions.

Enrichment for suppressors with increased sporulation at high salinity identifies mutations in kinA and sigH

On the basis of the results described above, we wondered whether it might be possible to isolate *B. subtilis* mutants that bypass the salt-imposed block in sporulation resulting in increased sporulation efficiency at high salinity. To this end, we used a genetic suppressor enrichment strategy. The wild-type *B. subtilis* strain JH642 was sporulated by nutrient exhaustion in DSM containing 1.2 M NaCl. After eliminating vegetative and sporulation-impaired cells by heat-kill (80°C for 20 min), serial dilutions were plated on LB-agar plates. The few heat-resistant colonies that emerged were then reinoculated into fresh DSM containing 1.2 M NaCl and the cycle was repeated. Mutants showing an increased sporulation phenotype in comparison to the JH642 wild type were isolated. Although the sporulation efficiencies of the 11 independently isolated suppressor strains were only moderately (threefold to eightfold) increased (Tables 1 and 2), we mapped the mutations in two of them by whole genome resequencing. Strikingly, in each

Table 1. Suppressor strains harbouring mutations in the *kinA* promoter region.

Strain	*kinA* promoter region sequence	Spores (%)	Fold increase
WT	TAGAAGGAGAATACTCATTTTCTAGCGAATCATACTAGGTAAAAGTCAATCTGTATATGTCGAAA - 14 nt – AAAGGAGGGATTCTGTG …	0.00016	-
NWB6	TAGAAGGAGAATACTCATTTTCTAGCGAATCATTCTAGGTAAAAGTCAATCTGTATATGTCGAAA - 14 nt – AAAGGAGGGATTCTGTG …	0.00086	5
NWB11	TAGAAGGAGATTACTCATTTTCTAGCGAATCATACTAGGTAAAAGTCAATCTGTATATGTCGAAA - 14 nt - AAAGGAGGGATTCTGTG …	0.00132	8
NWB13	TAGAAGGAGAAAACTCATTTTCTAGCGAATCATACTAGGTAAAAGTCAATCTGTATATGTCGAAA - 14 nt - AAAGGAGGGATTCTGTG …	0.00104	7
NWB16	TAGAAGGAGATTACTCATTTTCTAGCGAATCATACTAGTAAAAGTCAATCTGTATATGTCGAAA - 14 nt - AAAGGAGGGATTCTGTG …	0.00129	8
NWB17	TAGAAGGAGAATACTCATTTTCTA_CGAATCATACTAGGTAAAAGTCAATCTGTATATGTCGAAA - 14 nt - AAAGGAGGGATTCTGTG …	0.00078	5
NWB19	TAGAAGGAGAATACTCATTTTCTAGCGAATCATACTAGGTAAAAGTCAATCTGTATTTGTCGAAA - 14 nt - AAAGGAGGGATTCTGTG …	0.00097	6
NWB22	TAGAAGGAGAATACTCATTTTCTAGCGAATCAAACTAGGTAAAAGTCAATCTGTATATGTCGAAA - 14 nt - AAAGGAGGGATTCTGTG …	0.00065	4
NWB24	TAGAAGGAGAATACTCATTTTCTT_GCGAATCATACTAGGTAAAAGTCAATCTGTATATGTCGAAA - 14 nt - AAAGGAGGGATTCTGTG …	0.00081	5
NWB25	TAGAAGGA TAATACTCATTTTCTAGCGAATCATACTAGGTAAAAGTCAATCTGTATATGTCGAAA - 14 nt - AAAGGAGGGATTCTGTG …	0.00113	7

Suppressor strains that partially overcome the strong sporulation defect exhibited by the *B. subtilis* wild-type strain JH642 carry mutations in the promoter region of *kinA*. Mutations are highlighted in red. Predicted σ^H- (grey) and Spo0A- (black) recognition sites are indicated. Bold letters represent the start codon (GTG) of the *kinA* coding sequence and the ribosome-binding site is underlined. The spore titers and fold increases in sporulation efficiency of the suppressor strains are indicated.

Table 2. Suppressor strains harbouring mutations in the coding region of *sigH*.

Strain	Mutation	Spores (%)	Fold increase
WT	–	0.00016	–
NWB7	E146A [GAA→GCA]	0.00055	3
NWB21	I149F [ATT→TTT]	0.00067	4

Suppressor strains that partially overcome the strong sporulation defect exhibited by the *B. subtilis* wild-type strain JH642 carrying mutations in the coding region of *sigH*, the structural gene for the σ^H sporulation-specific sigma factor. The two mutations resulting in single amino acid substitutions in σ^H are indicated. The spore titers and fold increases in sporulation efficiency of the suppressor strains are indicated.

of these two suppressors only a single nucleotide polymorphism was observed in comparison to the JH642 reference genome (Smith *et al.*, 2014). In strain NWB6, we identified a point mutation in the promoter region of *kinA* (Table 1), and in strain NWB7 a point mutation in the coding region of the *sigH* gene was present that resulted in an amino acid substitution (E146A) in the σ^H protein (Table 2). With this information in hand, we PCR-amplified and sequenced the *kinA*, *sigH* and *spo0A* genes from the remaining nine suppressor strains. This analysis identified eight additional mutant alleles of the *kinA* promoter (Table 1) and one further allele of the *sigH* coding region (Table 2). None of the 11 suppressor strains harboured a mutation in *spo0A*. We also sequenced *kinA* and *sigH* (and their promoters) from *B. subtilis* 168, PY79, 3610 and DK1042 to investigate whether the strain-to-strain variation in sporulation efficiency at high salinity (Fig. 1C) could be explained by polymorphisms in these loci. Both genes and their promoter regions were identical to those in JH642, indicating that other differences among these strains account for the observed variability.

Overexpression of kinA upon entry into sporulation bypasses the early salt-stress imposed block

The identification of suppressor mutations in the promoter of *kinA* that modestly increased sporulation under high salinity (1.2 M NaCl) conditions, prompted us to investigate whether higher levels of KinA could overcome the salt-sensitive block. σ^H is required for the expression of both *spo0A* and *kinA* (see Fig. 2B) and thus it seemed possible that the overexpression of *kinA* could trigger enhanced sporulation in hypertonic medium. Accordingly, we used a strain in which *kinA* is expressed under the control of a strong IPTG-inducible promoter ($P_{hyperspank}$). Previous work from Fujita and coworkers has shown that increased expression of *kinA* using this expression system can bypass the need for starvation signals and induce sporulation during growth in rich medium

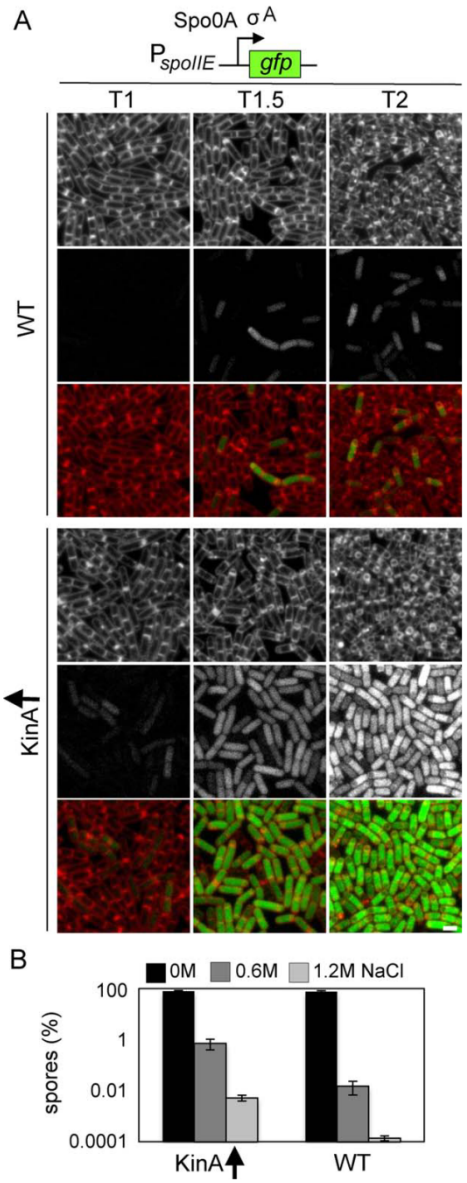


Fig. 6. Overexpression of *kinA* bypasses the early salt-sensitive block in sporulation.
A. Representative images of cells harbouring a Spo0A activity reporter (P_{spoIIE}-*gfp*) induced to sporulate in the presence of 0.6 M NaCl. Wild-type (WT; strain BDR2128) (Supporting Information Table S1) and a strain (BDR3087; Supporting Information Table S1) harbouring an IPTG-inducible allele of *kinA* were imaged at the indicated times after the initiation of sporulation. A 5 h sporulation time course of cells overexpressing *kinA* (BDR3087) comparing 0 M, 0.6 M and 1.2 M NaCl can be found in Supporting Information Fig. S9. Scale bar indicates 2 μm.
B. Bar graphs of sporulation efficiency in the wild-type (WT) and the *kinA* overexpression strain induced to sporulate in the presence of the indicated concentration of NaCl.

(Fujita and Losick, 2005; Eswaramoorthy *et al.*, 2009; Eswaramoorthy *et al.*, 2010). We induced sporulation in the presence of 0, 0.6 and 1.2 M NaCl and monitored Spo0A activity using the P_{spoIIE}-*gfp* reporter and followed sporulation with the fluorescent membrane dye TMA-DPH (Fig. 6A and Supporting Information Fig. S9A). Overexpression of *kinA* in all three conditions resulted in similar levels of KinA protein (Supporting Information Fig. S9B). Strikingly, KinA overproduction at 0.6 M NaCl resulted in high Spo0A activity and a statistically significant increase (>79%, $n > 430$) in the number of cells with polar divisions and engulfing forespores at hour 2 (T2) (Fig. 6A and Supporting Information Fig. S9). Even in the presence of 1.2 M NaCl, asymmetric septa and engulfing forespores were readily detectable, albeit with some delay compared to 0.6 M NaCl and the no-salt condition (Supporting Information Fig. S9). Thus, these results indicate that increasing the level of phosphorylated Spo0A can bypass the salt-sensitivity of σ^H and any other possible early blocks to sporulation.

Importantly, over-expressing *kinA* from the $P_{hyperspank}$ promoter significantly increased the number of heat-resistant spores in the presence of 0.6 M and 1.2 M NaCl (Fig. 6B). However, despite the near complete bypass of the early morphological events in 0.6 M NaCl (Fig. 6A and Supporting Information Fig. S9), the sporulation efficiency was still considerably reduced (by ~100-fold) compared to the no salt condition in which about 90% of the cells formed heat resistant spores (Fig. 6B). These data suggest that high salinity also impairs additional, later steps in the sporulation pathway.

High salinity impairs engulfment, σ^G activity and causes lysis

To define additional steps in the sporulation program that are sensitive to high salinity, we took advantage of a strain that harbours fluorescent reporters for all of the sporulation-specific sigma factors (Meeske *et al.*, 2016). In this strain, the reporter for the first compartment-specific transcription factor σ^F was a promoter (P_{spoIIQ}) fused to *yfp* (P_{spoIIQ}-*yfp*) (Londono-Vallejo *et al.*, 1997). The P_{spoIID} promoter that is recognized by the first

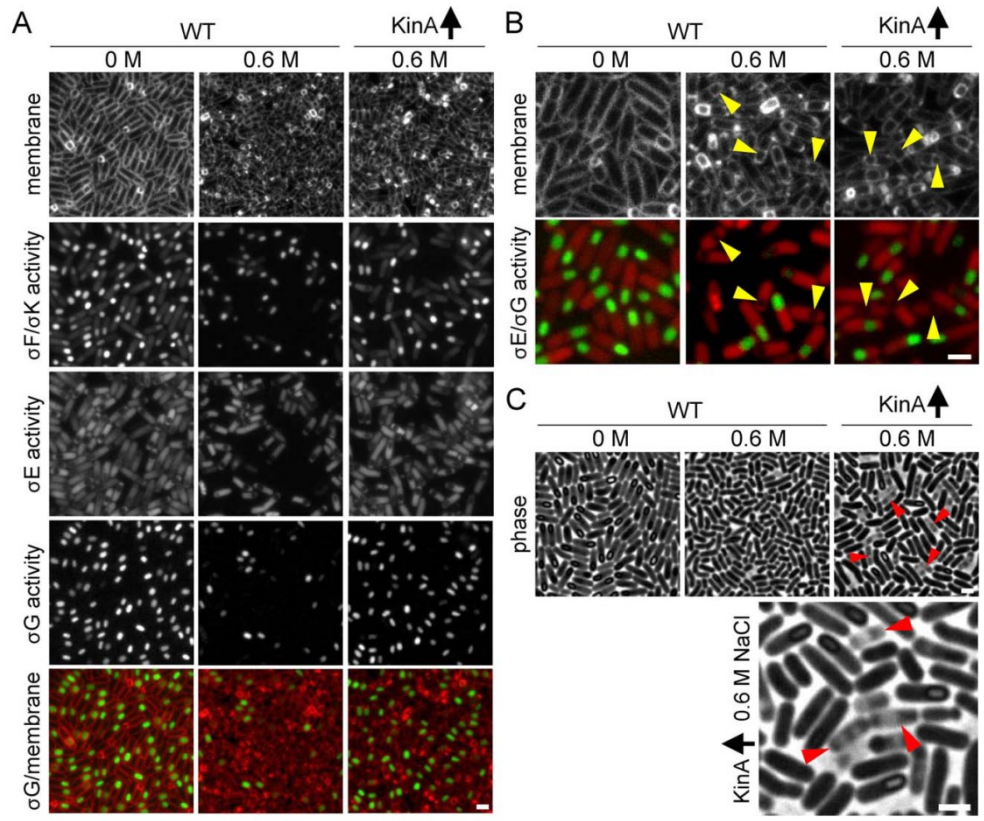


Fig. 7. High salinity impairs engulfment and late forespore gene expression and causes cell lysis.
A. Representative images of cells harbouring transcriptional reporters for all four stage- and compartment-specific sigma factors induced to sporulate in the presence of the indicated concentrations of NaCl. The membranes and reporters for σ^F (P_{spoIIQ}-*yfp*), σ^E (P_{spoIID}-mCherry), σ^G (P_{sspB}-*cfp*) and σ^K (P_{gerE}-*yfp*) are shown in black/white or as a merged image with membranes (red) and σ^G activity (green). Wild-type (WT) sporulating cells (strain BCR1071; Supporting Information Table S1) and a strain overexpressing *kinA* (BCR1274; Supporting Information Table S1) were imaged at hour 4 of sporulation. Additional time points can be found in Supporting Information Fig. S10A.
B. Representative images highlighting impaired engulfment and the defect in σ^G activation in a subset of sporulating cells grown in the presence of 0.6 M NaCl. Forespores with reduced σ^G activity are indicated (yellow arrow heads). Images show membranes (top) and an overlay of the σ^G activity reporter (green) and the σ^E activity reporter (red).
C. Representative phase contrast images from hour 5 of sporulation highlighting the increased frequency of cell lysis and reduction in phase-grey and phase-bright spores when cells bypass the early salt-sensitive block to sporulation. Additional time points can be found in Supporting Information Fig. S10B. Lysed cells are highlighted by red arrow heads. Scale bar indicates 2 μm.

mother-cell-specific transcription factor σ^E (Rong *et al.*, 1986) was fused to *mCherry* (P_{spoIID}-*mCherry*), and the reporter for the late forespore transcription factor σ^G was a promoter (P_{sspB}) fused to *cfp* (Nicholson *et al.*, 1989). Finally, the P_{gerE} promoter, recognized by the late mother-cell transcription factor σ^K (Cutting *et al.*, 1989), was fused to *yfp* (P_{gerE}-*yfp*). σ^F-dependent YFP fluorescence is first observed in the forespore at hour 2. By the time σ^K is activated in the mother cell (hour 4), σ^F activity is largely absent and YFP fluorescence is significantly reduced. Accordingly, we could use this fluorescent protein twice.

We introduced the $P_{hyperspank}$-*kinA* construct into this multireporter strain and induced sporulation in the presence of 0.6 M NaCl and IPTG. Over-expression of *kinA* resulted in activation of σ^F in the forespore and σ^E in the mother cell in almost all the sporulating cells (Fig. 7A and Supporting Information Fig. S10A). Interestingly,

many cells exhibited defects in the morphological process of engulfment: the migration of the mother-cell membranes around the forespore was uncoordinated and the engulfed forespores were smaller and contained membrane aberrations (Figs 6A, 7A and B, Supporting Information Figs S9A and S10A). Furthermore, although many sporulating cells displayed forespore-specific σ^G activity after the completion of engulfment, a subset of cells appeared to have reduced or undetectable levels (Fig. 7A and B and Supporting Information Fig. S10A). These phenotypes were also apparent in the 0.6 M NaCl condition where *kinA* was not overexpressed (Fig. 7A and B). Phase contrast images at later time points revealed a second and more striking phenotype: a reduction in phase grey spores and a concomitant increase in lysed cells (Fig. 7C and Supporting Information Fig. S10). Altogether, our results indicate that membrane remodeling and σ^G activity in the forespore display salt-sensitivity but the principal defect in spore formation in the presence of high salt is lysis of the sporulating cell prior to the maturation of the dormant spore.

Discussion

Building on previous observations (Kunst and Rapoport, 1995; Ruzal *et al.*, 1998), we report here that spore formation in *B. subtilis* is highly sensitive to even modest increases in the external salinity, an environmental challenge that leads to an almost exponential drop in sporulation efficiency once a threshold of 0.6 M NaCl is achieved (Fig. 1B). Interestingly, at this level of external salinity, growth of *B. subtilis* begins to be impaired (Boch *et al.*, 1994) and the cellular osmostress protection mechanisms of vegetative cells (e.g. the import of osmostress protectants and the synthesis of the compatible solutes) are already active (Brill *et al.*, 2011; Nau-Wagner *et al.*, 2012; Hoffmann *et al.*, 2013). Under extreme saline conditions (1.2 M NaCl) where growth of *B. subtilis* is more significantly retarded (Boch *et al.*, 1994), only one out of a million vegetative cells can form a mature, heat-resistant endospore (Fig. 1B). The goal of our study was to define the molecular basis for the salt-stress-imposed block in sporulation. Data reported by Sánchez-Rivas and coworkers (Ruzal and Sanchez-Rivas, 1998; Ruzal *et al.*, 1998) had already indicated that this happened early in the sporulation process but the underlying molecular mechanism(s) and the precise stage at which the salt-stress-imposed block occurred remained ill defined. Our data clarify these issues.

Our study shows that entry into sporulation by the *B. subtilis* cell is already impaired prior to polar division and the activation of the forespore-specific transcription factor σ^F (Fig. 2A, Supporting Information Figs S2 and S4). Since cells that have formed the asymmetric septum and activated σ^F are fully committed to the sporulation process (Errington, 2003; Dworkin and Losick, 2005; Higgins and Dworkin, 2012), the high-salinity-imposed block occurs prior to the 'point of no return'. Consistent with these findings, we further show that gene expression under the control of the master regulator of sporulation, Spo0A, and of the earliest acting sporulation-specific sigma factor σ^H, was strongly affected by high salinity (Fig. 3). Our data suggest that the molecular underpinnings for these defects rest on an insufficient cellular level of active Spo0A (Fig. 4), and the inability of σ^H to associate effectively with core RNA-polymerase (Fig. 5).

Previous work from Hecker and coworkers (Reder *et al.*, 2012a,b) demonstrated that the sigma factor (σ^B) that controls the general stress regulon (Hecker *et al.*, 2007; Price, 2011) activates *spo0E* encoding a phosphatase that attenuates flux through the phosphorelay. Although salt-shocks are among the strongest inducers of the general stress response in *B. subtilis* (Nannapaneni *et al.*, 2012), induction of σ^B-controlled genes by this environmental cue is only short-lived (Young *et al.*, 2013). Importantly, in cells experiencing chronic exposure to high salinity, as was used in our experiments, the general stress regulon is not induced (Spiegelhalter and Bremer, 1998). Accordingly, it is unlikely that the σ^B-dependent down-regulation of Spo0A~P makes a significant contribution to the early block in sporulation observed here. In support of this idea, the reduction in sporulation efficiency resulting from high salinity was indistinguishable between wild-type and an isogenic Δ*sigB* mutant (Supporting Information Fig. S11).

The genetic control of *spo0A* transcription is complex and involves just-in-time regulatory events (Chastanet *et al.*, 2010; Chastanet and Losick, 2011) that ensure that the cellular level of Spo0A~P increases in such a fashion that promoters with different affinities for this transcription factor become active at defined time points during starvation (Fujita and Losick, 2005; Fujita *et al.*, 2005). To initiate spore-formation, high threshold levels of Spo0A~P and proper activation dynamics are required (Grossman, 1995; Molle *et al.*, 2003; Fujita *et al.*, 2005; Vishnoi *et al.*, 2013). Our immunoblot analysis assessing Spo0A levels indicate that reduced amounts of Spo0A are present in cells cultured in the presence of 0.6 M NaCl and they are even further decreased when the salinity is increased to 1.2 M NaCl (Fig. 4). Hence, the inability of *B. subtilis* cells to sporulate efficiently at high salinity is likely explained by the insufficient accumulation of the master regulator of sporulation, Spo0A. Our data suggest that Spo0A protein is itself sensitive to high salinity (Figs 3B and 4) although

the molecular mechanism underlying this instability remains unknown. However, a second salt-sensitive event that contributes to the drop in Spo0A levels is the reduced association of σ^H with core RNA-polymerase (Fig. 5). Since the σ^H-dependent promoter of *spo0A* is a key genetic control element in setting cellular levels of Spo0A (Predich *et al.*, 1992; Chibazakura *et al.*, 1995; Eymann *et al.*, 2001), insufficient amounts of the σ^H-RNA-polymerase holoenzyme will contribute to a failure to achieve sporulation-promoting levels of Spo0A (Hoch, 1991; Predich *et al.*, 1992; Britton *et al.*, 2002).

We do not know the reasons why σ^H associates ineffectively with core RNA-polymerase (Fig. 5). One possibility is that an increase in ECF sigma factors in response to salt stress competes for core. However, we favour the idea that an altered composition of the cytoplasmic ion and solute pools is the culprit. Indeed, a *B. subtilis* mutant defective in the major Na^+-sodium extrusion system Mrp (Ito *et al.*, 1999; Swartz *et al.*, 2005) not only becomes highly sensitive to Na^+ ions but also exhibits a severe sporulation defect (Kosono *et al.*, 2000). *B. subtilis* typically maintains a very low cytoplasmic Na^+ level. As assessed by ^{23}Na-nuclear magnetic resonance spectroscopy, the intracellular Na^+ pool in an *mrp* mutant increases from practically nonmeasurable levels in the wild-type strain to approximately 12 mM under growth conditions where the external NaCl concentration was only 80 mM (Gorecki *et al.*, 2014). Since the Mrp system functions as a Na^+/H^+ antiporter, its operation is also connected to pH homeostasis (Swartz *et al.*, 2005; Krulwich *et al.*, 2011). It is, therefore, of interest to consider data reported by Zuber and coworkers who found, as reported here for salt-stressed cells, a defect in the association of σ^H with core RNA-polymerase in *B. subtilis* cells that were subjected to low pH stress (Liu *et al.*, 1999). The ion pool, in particular that of potassium, is certainly different in osmotically stressed cells from those that are not subjected to this challenge (Whatmore *et al.*, 1990). Furthermore, transient changes in the intracellular Na^+ content of high-salinity-exposed *B. subtilis* cells might occur as a consequence of the operation of osmotically induced and Na^+-driven osmostress protectant importers for glycine betaine (OpuD) and proline (OpuE) (Kappes *et al.*, 1996; von Blohn *et al.*, 1997).

Suppressor mutants with single nucleotide polymorphisms that map either in *sigH* or in the promoter region of *kinA* partially bypassed the salt-stress-imposed block in sporulation (Tables 1 and 2). Although these mutations only modestly enhanced sporulation under high salinity, they provided additional support for the idea that Spo0A and σ^H activities are impaired under these conditions. Consistent with our findings that the interaction between σ^H and core RNA polymerase is reduced in high salt, the two amino acid substitutions in σ^H (E146A and I149F) both reside in the putative helix-turn-helix located within subregion 3.1, which is associated with sigma binding to core RNA polymerase (Murakami and Darst, 2003). All the other suppressor mutations map in the regulatory region of *kinA* (Table 1), a gene that is dependent on σ^H and Spo0A for its expression (Predich *et al.*, 1992; Fujita and Sadaie, 1998a). Since these mutations likely increase *kinA* transcription, we tested a strain in which *kinA* could be overexpressed. Strikingly, KinA overproduction almost completely bypassed the block to entering the sporulation pathway in 0.6 M NaCl (Fig. 6A and Supporting Information Fig. S9). It also significantly increased the frequency of mature, heat-resistant spores; however, sporulation efficiency was still 100-fold lower than that observed in the absence of salt stress (Fig. 6B). This difference can be explained, in part, by impaired engulfment and reduced σ^G activity but is principally due to cell lysis prior to maturation of the spore (Fig. 7, Supporting Information Figs S9 and S10).

At first glance, the presence of a salt-stress imposed block to sporulation runs contrary to the idea that highly desiccation-resistant endospores (Setlow, 1995; Nicholson *et al.*, 2000; Setlow, 2014) might serve as an effective osmostress escape route when vegetative growth is no longer possible (Boch *et al.*, 1994). One possible explanation for this block is that *B. subtilis* has evolved an alternative strategy to survive starvation at high salinity that is more compatible with lower activities of σ^H and Spo0A. Indeed, we find that wild-type cells retain nearly complete viability in 1.2 M NaCl for at least 24 h after they have exhausted their nutrients (Supporting Information Fig. S12). After 24 h, the viability of these starved and salt-stressed cells steadily declines.

An alternative view places the strong and early block to sporulation by high salinity in a wider cellular context that takes the physiological status of starving (Errington, 2003; Eichenberger *et al.*, 2004; Higgins and Dworkin, 2012) and salt-stressed cells (Bremer and Krämer, 2000; Bremer, 2002) into account. The commitment to sporulate is a measure of last resort after the cell has run out of other options (Lopez *et al.*, 2009; Vlamakis *et al.*, 2013). It is a time-consuming and energy demanding process and involves the coordinated transcription of almost a quarter of the genome (Fawcett *et al.*, 2000; Molle *et al.*, 2003; Eichenberger *et al.*, 2004; Steil *et al.*, 2005; Nicolas *et al.*, 2012). Even under favourable laboratory conditions, it takes about 7 h to complete (Eichenberger *et al.*, 2004), and after polar division and the first compartment-specific transcription factors are activated, a point of no return in this developmental program is reached (Errington, 2003; Higgins and Dworkin, 2012). Hence, the *B. subtilis* cell

must assure that the dwindling nutritional and energetic resources available to it, and the environmental circumstances, will allow completion of the sporulation process. Otherwise, all things are lost: the mother cell will lyse and no spore will be formed.

High salinity environments impose considerable constraints on the physiology and growth of the *B. subtilis* cell (Boch *et al.*, 1994; Bremer and Krämer, 2000; Bremer, 2002). A key event in its defense against osmotic stress is the *de novo* synthesis and high-level accumulation of the compatible solute proline to maintain physiologically adequate levels of cellular hydration and turgor (Whatmore *et al.*, 1990; Brill *et al.*, 2011). The size of the proline pool is linearly dependent on the degree of the environmentally imposed osmotic stress and reaches about 500 mM in *B. subtilis* cells exposed to very high (1.2 M NaCl) saline growth conditions (Brill *et al.*, 2011; Hoffmann *et al.*, 2013; Zaprasis *et al.*, 2013). Genetic disruption of the osmostress-adaptive proline synthesis route makes *B. subtilis* highly salt sensitive, highlighting the important role of the pool of this compatible solute for growth under osmotically unfavourable circumstances (Brill *et al.*, 2011). The synthesis of a single proline molecule requires the expenditure of ~20 high-energy phosphate bonds (Akashi and Gojobori, 2002). Hence, the maintenance of proline up to levels of 500 mM (Hoffmann *et al.*, 2013; Zaprasis *et al.*, 2013) is an enormous physiological task for starving *B. subtilis* cells that are faced with entering the sporulation pathway while simultaneously experiencing severe osmotic stress. The data presented here suggest that the block to sporulation under salt stress is not an accident; rather we propose it serves as a safeguard to avert osmotically stressed *B. subtilis* cells from committing to a developmental program they cannot complete and would die trying.

Experimental procedures

General methods

B. subtilis strains were derived from 168, JH642 or PY79. Unless otherwise indicated, cells were grown in LB (Lennox) or CH (Sterlini–Mandelstam) media at 37°C. Sporulation was induced by resuspension at 37°C according to the method of Sterlini–Mandelstam or by nutrient exhaustion in supplemented DS medium (DSM complete) (Harwood and Cutting, 1990). Sporulation efficiency was determined in cultures grown for 24–36 h as the total number of heat-resistant (80°C for 20 min) colony-forming units (CFUs) compared with the total CFUs or heat-resistant CFUs of cells sporulated without NaCl. Tables of strains, plasmids and oligonucleotide primers and a description of strain and plasmid constructions can be found online as supplementary data (Supporting Information Tables S1, S2 and S3).

Fluorescence microscopy

Sporulating cells were concentrated by centrifugation at 8000 rpm for 1 min and immobilized on 2% agarose pads. Fluorescence microscopy was performed using an Olympus BX61 microscope equipped with a UplanF1 100X phase contrast objective lens and a CoolSnapHQ digital camera (Photometrics) or a Nikon TE2000 inverted microscope with a Nikon CFI Plan Apo VC 100X objective lens. Images were acquired using Metamorph software (Molecular DEVICES, Sunnyvale, CA, USA). Membranes were stained with TMA-DPH (50 μM) (Molecular Probes; ThermoFisher Scientific) and fission of mother cell membranes was assessed as previously described (Doan *et al.*, 2013). Image analysis and processing was performed in Metamorph.

Immunoblot analysis

Whole-cell lysates from sporulating cells were prepared as previously described (Doan and Rudner, 2007). Samples were heated for 5 min at 65°C prior to loading. Equivalent loading of proteins was based on the OD_{600} of the cell cultures at the time of harvest. Samples were separated on a 12.5% SDS-polyacrylamide gel and transferred to a methanol-activated PVDF membrane. Membranes were blocked in 5% nonfat milk with 0.5% Tween-20 for 1 h. Blocked membranes were probed with anti-σ^H (diluted 1:2500), anti-Spo0A (diluted 1:5000) or anti-SpoIIQ (diluted 1:10,000) (Doan *et al.*, 2009), anti-σ^A (diluted 1:10,000), anti-KinA (diluted 1:5000) and anti-core RNA-polymerase (diluted 1:5000). These primary antibodies were diluted into PBS with 0.05% Tween-20 and incubations were carried out at 4°C overnight. Primary antibodies were detected with horseradish-peroxidase conjugated anti-mouse or anti-rabbit antibodies and detected with Western Lightning ECL reagent as described by the manufacturer (PerkinElmer, Waltham, MA, USA).

Protein pull-down assay

25-ml cultures of strain BDR485 (Supporting Information Table S1) harbouring a functional *rpoC-his6* fusion were grown in CH medium in the presence and absence of 0.6 M NaCl. At an OD_{600} of ~0.5, sporulation was induced by resuspension in the presence and absence of 0.6 M NaCl. 1.5 hours later, 20 ml from each culture were harvested by centrifugation (5000 rpm for 10 min at room temperature), washed two times in 10 ml cold lysis buffer (20 mM Tris pH7.5, 300 mM NaCl, 20 mM imidazole, 5 mM $MgCl_2$, 5 mM β-mercaptoethanol, 1 mM PMSF, 1 μg ml^{-1} leupeptin and 1 μg ml^{-1} pepstatin) and the pellets were flash frozen and stored at −80°C. The cell pellets were thawed and resuspended in 1 ml cold lysis buffer with 2 mg/ml lysozyme and incubated at 4°C for 20 min. The cell suspension was lysed by sonication followed by the addition of 10 μl of 1 mg/ml DNase (Worthington, Lakewood, NJ, USA), 10 μl of 10 mg/ml RNaseA and 20 μl of 100 mM PMSF. The lysate was clarified by centrifugation at 13,000 rpm at 4°C for 5 min and the supernatant was

incubated with 50 µl of Ni^{2+}-agarose (Qiagen Hilden, Germany) for 1 h at 4°C. The resin was washed 5 times with 1.5 ml wash buffer (20 mM Tris pH 7.5, 300 mM NaCl, 20 mM imidazole and 5 mM β-mercaptoethanol) and then resuspended in 100 µl SDS sample buffer and heated at 80°C for 10 min. Protein fractions were separated on a 12.5% SDS-polyacrylamide gel and visualized with Coomassie brilliant blue or by immunoblot.

Suppressor enrichment

To enrich for mutants that could form heat-resistant spores in the presence of high salinity, cells of the wild-type *B. subtilis* strain JH642 (Smith *et al.*, 2014) were sporulated (30 h after reaching stationary phase) in DS medium containing 1.2 M NaCl. After eliminating vegetative cells by pasteurization (80°C, 20 min), serial dilutions were plated onto LB agar plates and the spore titer was determined. Colonies arising from spores were then reinoculated in fresh DS medium containing 1.2 M NaCl. After 30 h of growth after reaching stationary phase, the spore titer of the cultures was again determined. Cultures that showed an increased spore titer in comparison to that of the JH642 wild-type strain were analyzed for mutations either by whole genome resequencing, or by targeted sequencing of PCR products derived from the *spo0A*, *kinA* and *sigH* genes. The DNA primers used to amplify these genes and their flanking regions can be found in Supporting Information Table S3.

Mapping suppressors by whole genome sequencing and PCR analysis

Chromosomal DNA from *B. subtilis* was isolated using the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany). To identify the mutation(s) in the suppressor mutant strains NWB6 and NWB7 (Table 1), the genomic DNA was subjected to whole-genome sequencing. Library preparation and sequencing analysis were performed by the Göttingen Genomics Laboratory (Göttingen, Germany). The reads were mapped on the reference genome of the *B. subtilis* JH642 strain (GenBank accession number: CP_007800) (Smith *et al.*, 2014). Mapping of the reads was performed as previously described (Zaprasis *et al.*, 2014) using the Geneious software package (Biomatters Ltd., New Zealand). Single nucleotide polymorphisms (SNPs) were considered as significant when the total coverage depth exceeded 25 reads with a frequency variance of >90%. For the molecular analysis of all other suppressor strains, the promoter and coding regions of *kinA*, *sigH* and *spo0A* were PCR-amplified and Sanger sequenced (MWG Eurofins, Ebersberg, Germany).

Acknowledgements

This work was supported by contributions of the Fonds der Chemischen Industrie (to F.M.C and E.B.), the Max-Buchner-Forschungsstiftung (MBFSt-Kennziffer 3381) (to F.M.C) and the National Institute of Health grant GM086466 to D.Z.R. We thank Masaya Fujita (University of Houston, TX, USA) for generously providing antibodies and an anonymous reviewer for noting the location of the amino acid substitutions in σ^H. Jörg Schuldes and Rolf Daniel from the Göttingen Genomics Laboratory (G2L) (Göttingen, Germany) are acknowledged for kindly performing genome re-sequencing. We appreciate the expert help of Vickie Koogle in the editing of the manuscript.

N.W. and K.E.F. are recipients of Ph.D. fellowships from the International Max-Planck Research School for Environmental, Cellular and Molecular Microbiology (IMPRS-Mic, Marburg) and gratefully acknowledge its generous support. N.W. thankfully acknowledges the receipt of an EMBO short-term fellowship that funded his stay in the laboratory of D.Z.R at the Harvard Medical School (Boston, MA, USA). F.G. R-G. is a recipient of a Conacyt postdoctoral fellowship (Mexico).

References

Akashi, H., and Gojobori, T. (2002) Metabolic efficiency and amino acid composition in the proteomes of *Escherichia coli* and *Bacillus subtilis*. *Proc Natl Acad Sci USA* **99**: 3695–3700.

Bashir, A., Hoffmann, T., Kempf, B., Xie, X., Smits, S.H., and Bremer, E. (2014) Plant-derived compatible solutes proline betaine and betonicine confer enhanced osmotic and temperature stress tolerance to *Bacillus subtilis*. *Microbiology* **160**: 2283–2294.

Belda, E., Sekowska, A., Le Fevre, F., Morgat, A., Mornico, D., Ouzounis, C., *et al.* (2013) An updated metabolic view of the *Bacillus subtilis* 168 genome. *Microbiology* **159**: 757–770.

von Blohn, C., Kempf, B., Kappes, R.M., and Bremer, E. (1997) Osmostress response in *Bacillus subtilis*: characterization of a proline uptake system (OpuE) regulated by high osmolarity and the alternative transcription factor sigma B. *Mol Microbiol* **25**: 175–187.

Boch, J., Kempf, B., and Bremer, E. (1994) Osmoregulation in *Bacillus subtilis*: synthesis of the osmoprotectant glycine betaine from exogenously provided choline. *J Bacteriol* **176**: 5364–5371.

Bremer, E. (2002) Adaptation to changing osmolarity. In *Bacillus subtilis and its Closes Relatives: From Genes to Cells*. Sonenshein, A.L., Hoch, J.A., and Losick, R. (eds). Washington, DC: ASM Press, pp. 385–391.

Bremer, E., and Krämer, R. (2000) Coping with osmotic challenges: osmoregulation through accumulation and release of compatible solutes. In *Bacterial Stress Responses*. Storz, G., and Hengge-Aronis, R. (eds). Washington, DC: ASM Press, pp. 79–97.

Brill, J., Hoffmann, T., Bleisteiner, M., and Bremer, E. (2011) Osmotically controlled synthesis of the compatible solute proline is critical for cellular defense of *Bacillus subtilis* against high osmolarity. *J Bacteriol* **193**: 5335–5346.

Britton, R.A., Eichenberger, P., Gonzalez-Pastor, J.E., Fawcett, P., Monson, R., Losick, R., and Grossman, A.D. (2002) Genome-wide analysis of the stationary-phase sigma factor (sigma-H) regulon of *Bacillus subtilis*. *J Bacteriol* **184**: 4881–4890.

Broy, S., Chen, C., Hoffmann, T., Brock, N.L., Nau-Wagner, G., Jebbar, M., *et al.* (2015) Abiotic stress protection by

ecologically abundant dimethylsulfoniopropionate and its natural and synthetic derivatives: insights from *Bacillus subtilis*. *Environ Microbiol* **7**: 2362–2378.
Burbulys, D., Trach, K.A., and Hoch, J.A. (1991) Initiation of sporulation in *Bacillus subtilis* is controlled by a multicomponent phosphorelay. *Cell* **64**: 545–552.
Chastanet, A., and Losick, R. (2011) Just-in-time control of Spo0A synthesis in *Bacillus subtilis* by multiple regulatory mechanisms. *J Bacteriol* **193**: 6366–6374.
Chastanet, A., Vitkup, D., Yuan, G.C., Norman, T.M., Liu, J.S., and Losick, R.M. (2010) Broadly heterogeneous activation of the master regulator for sporulation in *Bacillus subtilis*. *Proc Natl Acad Sci USA* **107**: 8486–8491.
Chibazakura, T., Kawamura, F., Asai, K., and Takahashi, H. (1995) Effects of *spo0* mutations on *spo0A* promoter switching at the initiation of sporulation in *Bacillus subtilis*. *J Bacteriol* **177**: 4520–4523.
Chung, J.D., Stephanopoulos, G., Ireton, K., and Grossman, A.D. (1994) Gene expression in single cells of *Bacillus subtilis*: evidence that a threshold mechanism controls the initiation of sporulation. *J Bacteriol* **176**: 1977–1984.
Cutting, S., Panzer, S., and Losick, R. (1989) Regulatory studies on the promoter for a gene governing synthesis and assembly of the spore coat in *Bacillus subtilis*. *J Mol Biol* **207**: 393–404.
Doan, T., and Rudner, D.Z. (2007) Perturbations to engulfment trigger a degradative response that prevents cell-cell signalling during sporulation in *Bacillus subtilis*. *Mol Microbiol* **64**: 500–511.
Doan, T., Morlot, C., Meisner, J., Serrano, M., Henriques, A.O., Moran, C.P., Jr, and Rudner, D.Z. (2009) Novel secretion apparatus maintains spore integrity and developmental gene expression in *Bacillus subtilis*. *PLoS Genet* **5**: e1000566.
Doan, T., Coleman, J., Marquis, K.A., Meeske, A.J., Burton, B.M., Karatekin, E., and Rudner, D.Z. (2013) FisB mediates membrane fission during sporulation in *Bacillus subtilis*. *Genes Dev* **27**: 322–334.
Dworkin, J., and Losick, R. (2005) Developmental commitment in a bacterium. *Cell* **121**: 401–409.
Earl, A.M., Losick, R., and Kolter, R. (2008) Ecology and genomics of *Bacillus subtilis*. *Trends Microbiol* **16**: 269–275.
Earl, A.M., Eppinger, M., Fricke, W.F., Rosovitz, M.J., Rasko, D.A., Daugherty, S., *et al.* (2012) Whole-genome sequences of *Bacillus subtilis* and close relatives. *J Bacteriol* **194**: 2378–2379.
Eichenberger, P., Fujita, M., Jensen, S.T., Conlon, E.M., Rudner, D.Z., Wang, S.T., *et al.* (2004) The program of gene transcription for a single differentiating cell type during sporulation in *Bacillus subtilis*. *PLoS Biol* **2**: e328.
Errington, J. (2003) Regulation of endospore formation in *Bacillus subtilis*. *Nat Rev Microbiol* **1**: 117–126.
Eswaramoorthy, P., Guo, T., and Fujita, M. (2009) In vivo domain-based functional analysis of the major sporulation sensor kinase, KinA, in *Bacillus subtilis*. *J Bacteriol* **191**: 5358–5368.
Eswaramoorthy, P., Duan, D., Dinh, J., Dravis, A., Devi, S.N., and Fujita, M. (2010) The threshold level of the sensor histidine kinase KinA governs entry into sporulation in *Bacillus subtilis*. *J Bacteriol* **192**: 3870–3882.
Eymann, C., Mittenhuber, G., and Hecker, M. (2001) The stringent response, sigmaH-dependent gene expression and sporulation in *Bacillus subtilis*. *Mol Gen Genet* **264**: 913–923.
Fawcett, P., Eichenberger, P., Losick, R., and Youngman, P. (2000) The transcriptional profile of early to middle sporulation in *Bacillus subtilis*. *Proc Natl Acad Sci USA* **97**: 8063–8068.
Fujita, M., and Losick, R. (2003) The master regulator for entry into sporulation in *Bacillus subtilis* becomes a cell-specific transcription factor after asymmetric division. *Genes Dev* **17**: 1166–1174.
Fujita, M., and Losick, R. (2005) Evidence that entry into sporulation in *Bacillus subtilis* is governed by a gradual increase in the level and activity of the master regulator Spo0A. *Genes Dev* **19**: 2236–2244.
Fujita, M., and Sadaie, Y. (1998a) Feedback loops involving Spo0A and AbrB in in vitro transcription of the genes involved in the initiation of sporulation in *Bacillus subtilis*. *J Biochem* **124**: 98–104.
Fujita, M., and Sadaie, Y. (1998b) Promoter selectivity of the *Bacillus subtilis* RNA polymerase sigmaA and sigmaH holoenzymes. *J Biochem* **124**: 89–97.
Fujita, M., and Sadaie, Y. (1998c) Rapid isolation of RNA polymerase from sporulating cells of *Bacillus subtilis*. *Gene* **221**: 185–190.
Fujita, M., Gonzalez-Pastor, J.E., and Losick, R. (2005) High- and low-threshold genes in the Spo0A regulon of *Bacillus subtilis*. *J Bacteriol* **187**: 1357–1368.
Gorecki, K., Hagerhall, C., and Drakenberg, T. (2014) The Na+ transport in gram-positive bacteria defect in the Mrp antiporter complex measured with ^{23}Na nuclear magnetic resonance. *Anal Biochem* **445**: 80–86.
Grossman, A.D. (1995) Genetic networks controlling the initiation of sporulation and the development of genetic competence in *Bacillus subtilis*. *Annu Rev Genet* **29**: 477–508.
Haldenwang, W.G. (1995) The sigma factors of *Bacillus subtilis*. *Microbiol Rev* **59**: 1–30.
Harwood, C.R., and Cutting, S.M. (1990) *Molecular Biological Methods for Bacillus*. Chichester: Wiley.
Hecker, M., Pane-Farre, J., and Völker, U. (2007) SigB-dependent general stress response in *Bacillus subtilis* and related gram-positive bacteria. *Annu Rev Microbiol* **61**: 215–236.
Higgins, D., and Dworkin, J. (2012) Recent progress in *Bacillus subtilis* sporulation. *FEMS Microbiol Rev* **36**: 131–148.
Hoch, J.A. (1991) *spo0* genes, the phosphorelay, and the initiation of sporulation. In *Bacillus subtilis and Other Grampositive Bacteria: Biochemistry, Physiology, and Molecular Genetics*. Sonenshein, A.L., Hoch, J.A., and Losick, R. (eds). Washington, DC: ASM Press, pp. 747–755.
Hoch, J.A. (1993) Regulation of the phosphorelay and the initiation of sporulation in *Bacillus subtilis*. *Annu Rev Microbiol* **47**: 441–465.
Hoffmann, T., Wensing, A., Brosius, M., Steil, L., Volker, U., and Bremer, E. (2013) Osmotic control of *opuA* expression in *Bacillus subtilis* and its modulation in response to intracellular glycine betaine and proline pools. *J Bacteriol* **195**: 510–522.

Holtmann, G., Bakker, E.P., Uozumi, N., and Bremer, E. (2003) KtrAB and KtrCD: two K^+ uptake systems in *Bacillus subtilis* and their role in adaptation to hypertonicity. *J Bacteriol* **185**: 1289–1298.

Ito, M., Guffanti, A.A., Oudega, B., and Krulwich, T.A. (1999) *mrp*, a multigene, multifunctional locus in *Bacillus subtilis* with roles in resistance to cholate and to Na^+ and in pH homeostasis. *J Bacteriol* **181**: 2394–2402.

Kappes, R.M., Kempf, B., and Bremer, E. (1996) Three transport systems for the osmoprotectant glycine betaine operate in *Bacillus subtilis*: characterization of OpuD. *J Bacteriol* **178**: 5071–5079.

Kosono, S., Ohashi, Y., Kawamura, F., Kitada, M., and Kudo, T. (2000) Function of a principal $Na^{(+)}/H^{(+)}$ antiporter, ShaA, is required for initiation of sporulation in *Bacillus subtilis*. *J Bacteriol* **182**: 898–904.

Krulwich, T.A., Sachs, G., and Padan, E. (2011) Molecular aspects of bacterial pH sensing and homeostasis. *Nat Rev Microbiol* **9**: 330–343.

Kunst, F., and Rapoport, G. (1995) Salt stress is an environmental signal affecting degradative enzyme synthesis in *Bacillus subtilis*. *J Bacteriol* **177**: 2403–2407.

Liu, J., and Zuber, P. (2000) The ClpX protein of *Bacillus subtilis* indirectly influences RNA polymerase holoenzyme composition and directly stimulates sigma-dependent transcription. *Mol Microbiol* **37**: 885–897.

Liu, J., Cosby, W.M., and Zuber, P. (1999) Role of lon and ClpX in the post-translational regulation of a sigma subunit of RNA polymerase required for cellular differentiation in *Bacillus subtilis*. *Mol Microbiol* **33**: 415–428.

Londono-Vallejo, J.A., Frehel, C., and Stragier, P. (1997) SpoIIQ, a forespore-expressed gene required for engulfment in *Bacillus subtilis*. *Mol Microbiol* **24**: 29–39.

Lopez, D., Vlamakis, H., and Kolter, R. (2009) Generation of multiple cell types in *Bacillus subtilis*. *FEMS Microbiol Rev* **33**: 152–163.

Mandic-Mulec, I., Stefanic, P., and van Elsas, J.D. (2015) Ecology of Bacillaceae. *Microbiol Spectr* **3**(2): TBS-0017-2013.

Meeske, A.J., Rodrigues, C.D.A., Brady, J., Lim, H.C., Benhardt, T.G., and Rudner, D.Z. (2016) High throughput genetic screens identify a large and diverse collection of new sporulation genes in *Bacillus sublitils*. *PLoS Biol* **14**(1): e1002341.

Molle, V., Fujita, M., Jensen, S.T., Eichenberger, P., Gonzalez-Pastor, J.E., Liu, J.S., and Losick, R. (2003) The Spo0A regulon of *Bacillus subtilis*. *Mol Microbiol* **50**: 1683–1701.

Murakami, K.S., and Darst, S.A. (2003) Bacterial RNA polymerases: the whol story. *Curr Opin Struct Biol* **13**: 31–39.

Nannapaneni, P., Hertwig, F., Depke, M., Hecker, M., Mäder, U., Völker, U., *et al*. (2012) Defining the structure of the general stress regulon of *Bacillus subtilis* using targeted microarray analysis and random forest classification. *Microbiology* **158**: 696–707.

Nau-Wagner, G., Opper, D., Rolbetzki, A., Boch, J., Kempf, B., Hoffmann, T., and Bremer, E. (2012) Genetic control of osmoadaptive glycine betaine synthesis in *Bacillus subtilis* through the choline-sensing and glycine betaine-responsive GbsR repressor. *J Bacteriol* **194**: 2703–2714.

Nicholson, W.L., Sun, D.X., Setlow, B., and Setlow, P. (1989) Promoter specificity of sigma G-containing RNA polymerase from sporulating cells of *Bacillus subtilis*: identification of a group of forespore-specific promoters. *J Bacteriol* **171**: 2708–2718.

Nicholson, W.L., Munakata, N., Horneck, G., Melosh, H.J., and Setlow, P. (2000) Resistance of *Bacillus* endospores to extreme terrestrial and extraterrestrial environments. *Microbiol Mol Biol Rev* **64**: 548–572.

Nicolas, P., Mader, U., Dervyn, E., Rochat, T., Leduc, A., Pigeonneau, N., *et al*. (2012) Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science* **335**: 1103–1106.

Phillips, Z.E., and Strauch, M.A. (2002) *Bacillus subtilis* sporulation and stationary phase gene expression. *Cell Mol Life Sci* **59**: 392–402.

Predich, M., Nair, G., and Smith, I. (1992) *Bacillus subtilis* early sporulation genes *kinA*, *spo0F*, and *spo0A* are transcribed by the RNA polymerase containing sigma H. *J Bacteriol* **174**: 2771–2778.

Price, C.W. (2011) General stress responses in *Bacillus subtilis* and related Gram-positive bacteria. In *Bacterial Stress Responses*. Storz, G., and Hengge, R. (eds). Washington, DC: ASM Press, pp. 301–318.

Reder, A., Albrecht, A., Gerth, U., and Hecker, M. (2012a) Cross-talk between the general stress response and sporulation initiation in *Bacillus subtilis* – the σB promoter of *spo0E* represents an AND-gate. *Mol Microbiol* **10**: 2741–2756.

Reder, A., Gerth, U., and Hecker, M. (2012b) Integration of σB activity into the decision-making process of sporulation initiation in *Bacillus subtilis*. *J Bacteriol* **194**: 1065–1074.

Rodrigues, C.D., Marquis, K.A., Meisner, J., and Rudner, D.Z. (2013) Peptidoglycan hydrolysis is required for assembly and activity of the transenvelope secretion complex during sporulation in *Bacillus subtilis*. *Mol Microbiol* **89**: 1039–1052.

Rong, S., Rosenkrantz, M.S., and Sonenshein, A.L. (1986) Transcriptional control of the *Bacillus subtilis spoIID* gene. *J Bacteriol* **165**: 771–779.

Ruzal, S.M., and Sanchez-Rivas, C. (1998) In *Bacillus subtilis* DegU-P is a positive regulator of the osmotic response. *Curr Microbiol* **37**: 368–372.

Ruzal, S.M., Lopez, C., Rivas, E., and Sanchez-Rivas, C. (1998) Osmotic strength blocks sporulation at stage II by impeding activation of early sigma factors in *Bacillus subtilis*. *Curr Microbiol* **36**: 75–79.

Segall, J., and Losick, R. (1977) Cloned *Bacillus subtilis* DNA containing a gene that is activated early during sporulation. *Cell* **11**: 751–761.

Setlow, P. (1995) Mechanisms for the prevention of damage to DNA in spores of *Bacillus* species. *Annu Rev Microbiol* **49**: 29–54.

Setlow, P. (2014) Germination of spores of *Bacillus* species: what we know and do not know. *J Bacteriol* **196**: 1297–1305.

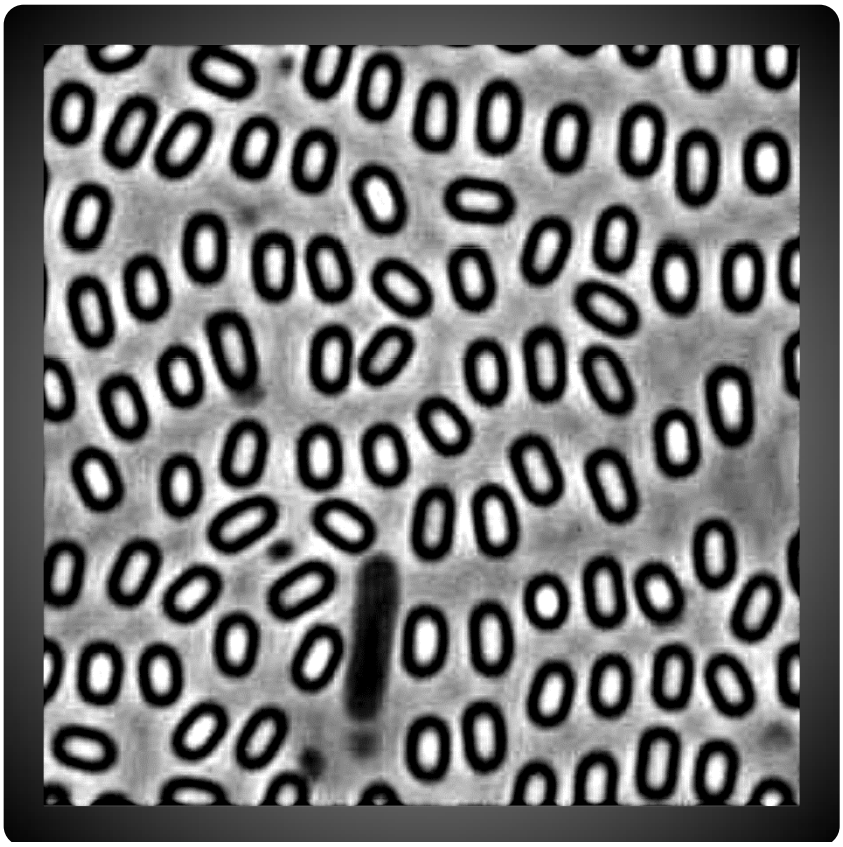
Sinai, L., Rosenberg, A., Smith, Y., Segev, E., and Ben-Yehuda, S. (2015) The molecular timeline of a reviving bacterial spore. *Mol Cell* **57**: 695–707.

Smith, J.L., Goldberg, J.M., and Grossman, A.D. (2014) Complete genome sequences of *Bacillus subtilis* subsp. subtilis laboratory strains JH642 (AG174) and AG1839. *Genome Announc* **2**: e00663– e00714.

Sonenshein, A.L. (2000) Control of sporulation initiation in *Bacillus subtilis*. *Curr Opin Microbiol* **3**: 561–566.
Sonenshein, A.L., Hoch, J.A., and Losick, R. (2002) *Bacillus subtilis and its Closest Relatives. From Genes to Cells*. Washingtin, DC: ASM Press.
Spiegelhalter, F., and Bremer, E. (1998) Osmoregulation of the *opuE* proline transport gene from *Bacillus subtilis*: contributions of the sigma A- and sigma B-dependent stress-responsive promoters. *Mol Microbiol* **29**: 285–296.
Steil, L., Serrano, M., Henriques, A.O., and Völker, U. (2005) Genome-wide analysis of temporally regulated and compartment-specific gene expression in sporulating cells of *Bacillus subtilis*. *Microbiology* **151**: 399–420.
Stragier, P., and Losick, R. (1996) Molecular genetics of sporulation in *Bacillus subtilis*. *Annu Rev Genet* **30**: 297–341.
Strauch, M., Webb, V., Spiegelman, G., and Hoch, J.A. (1990) The SpoOA protein of *Bacillus subtilis* is a repressor of the *abrB* gene. *Proc Natl Acad Sci USA* **87**: 1801–1805.
Strauch, M.A., Trach, K.A., Day, J., and Hoch, J.A. (1992) Spo0A activates and represses its own synthesis by binding at its dual promoters. *Biochimie* **74**: 619–626.
Sturm, A., and Dworkin, J. (2015) Phenotypic diversity as a mechanism to exit cellular dormany. *Curr Biol* **25**: 2272–2277.
Swartz, T.H., Ikewada, S., Ishikawa, O., Ito, M., and Krulwich, T.A. (2005) The Mrp system: a giant among monovalent cation/proton antiporters? *Extremophiles* **9**: 345–354.
Vishnoi, M., Narula, J., Devi, S.N., Dao, H.A., Igoshin, O.A., and Fujita, M. (2013) Triggering sporulation in *Bacillus subtilis* with artificial two-component systems reveals the importance of proper Spo0A activation dynamics. *Mol Microbiol* **90**: 181–194.
Vlamakis, H., Chai, Y., Beauregard, P., Losick, R., and Kolter, R. (2013) Sticking together: building a biofilm the *Bacillus subtilis* way. *Nat Rev Microbiol* **11**: 157–168.
Whatmore, A.M., Chudek, J.A., and Reed, R.H. (1990) The effects of osmotic upshock on the intracellular solute pools of *Bacillus subtilis*. *J Gen Microbiol* **136**: 2527–2535.
Wood, J.M. (2011) Bacterial osmoregulation: a paradigm for the study of cellular homeostasis. *Annu Rev Microbiol* **65**: 215–238.
York, K., Kenney, T.J., Satola, S., Moran, C.P., Jr, Poth, H., and Youngman, P. (1992) Spo0A controls the sigma A-dependent activation of *Bacillus subtilis* sporulation-specific transcription unit *spoIIE*. *J Bacteriol* **174**: 2648–2658.
Young, J.W., Locke, J.C., and Elowitz, M.B. (2013) Rate of environmental change determines stress response specificity. *Proc Natl Acad Sci USA* **110**: 4140–4145.
Zaprasis, A., Brill, J., Thüring, M., Wünsche, G., Heun, M., Barzantny, H., *et al.* (2013) Osmoprotection of *Bacillus subtilis* through import and proteolysis of proline-containing peptides. *Appl Environ Microbiol* **79**: 567–587.
Zaprasis, A., Hoffmann, T., Stannek, L., Gunka, K., Commichau, F.M., and Bremer, E. (2014) The gamma-aminobutyrate permease GabP serves as the third proline transporter of *Bacillus subtilis*. *J Bacteriol* **196**: 515–526.
Zuber, P., and Losick, R. (1987) Role of AbrB in Spo0A- and Spo0B-dependent utilization of a sporulation promoter in *Bacillus subtilis*. *J Bacteriol* **169**: 2223–2230.

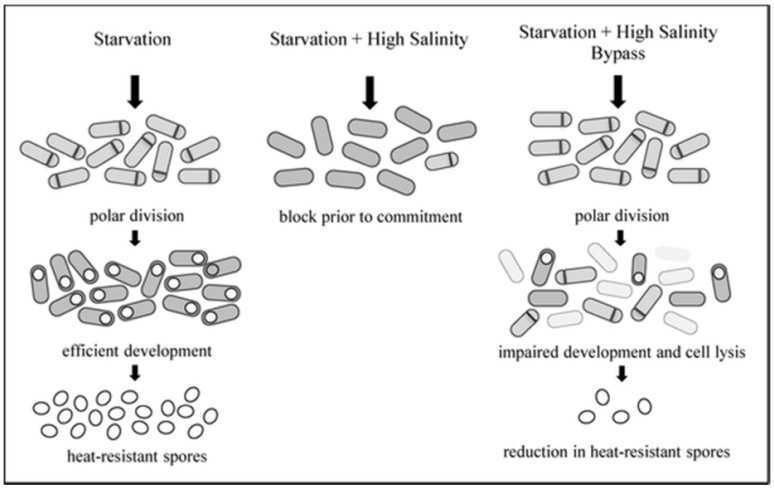
Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site.



Cover Illustration:

Image of endospores derived from *B. subtilis* cells and a single vegetative cell.
Submitted as ‘Cover Illustration‘ to *Molecular Microbiology* in connection with the following manuscript (*in press*).



Widderich & Rodrigues and co-workers investigated how Bacillus subtilis simultaneously responds to two common challenges that it frequently encounters in its native soil habitat: starvation and osmotic stress. At high salinity, starving B. subtilis cells do not enter the sporulation pathway, failing to activate sporulation genes under the control of the master regulators Spo0A and SigH. In so doing, they prevent osmotically stressed cells from committing to a developmental program they cannot complete.
50x31mm (300 x 300 DPI)

Salt-sensitivity of SigH and Spo0A prevents *Bacillus subtilis* sporulation at high osmolarity avoiding death during cellular differentiation

Nils Widderich[1,¶], Christopher D.A. Rodrigues[2,¶], Fabian M. Commichau[3], Kathleen E. Fischer[1], Fernando H. Ramirez-Guadiana[2], David Z. Rudner[2,*] and Erhard Bremer[1,*]

[1]Department of Biology, Laboratory for Molecular Microbiology, Philipps-University Marburg, Karl-von-Frisch Str. 8, D-35043 Marburg, Germany

[2]Harvard Medical School, Department of Microbiology and Immunobiology, 77 Avenue Louis Pasteur, Boston, MA 02115-5701, USA

[3]Department of General Microbiology, Institute of Microbiology and Genetics, Georg August University Göttingen, Griesebachstr. 8, D-37077 Göttingen, Germany

Running title: High salinity blocks sporulation

Keywords: salt stress – development – sporulation – sigma factors – kinase – cell death

For correspondence during the reviewing and editorial process:

Erhard Bremer: Philipps-University Marburg, Dept. of Biology, Laboratory for Microbiology, Karl-von-Frisch-Str. 8, D-35032 Marburg, Germany. Phone: (+49)-6421-2821529. Fax: (+49)-6421-2828979. E-Mail: bremer@staff.uni-marburg.de

---

[¶]These authors contributed equally to this work

*For correspondence. E.B.: E-mail bremer@staff.uni-marburg.de; Tel. (+49)-6421-2821529; (+49)-6421-2828979; D.Z.R.: rudner@hms.harvard.edu; (+1)-617-432-4455; (+1)-617-432-4787

SUPPLEMENTARY MATERIAL

Supplemental Figures

Widderich *et al.*, Figure S1

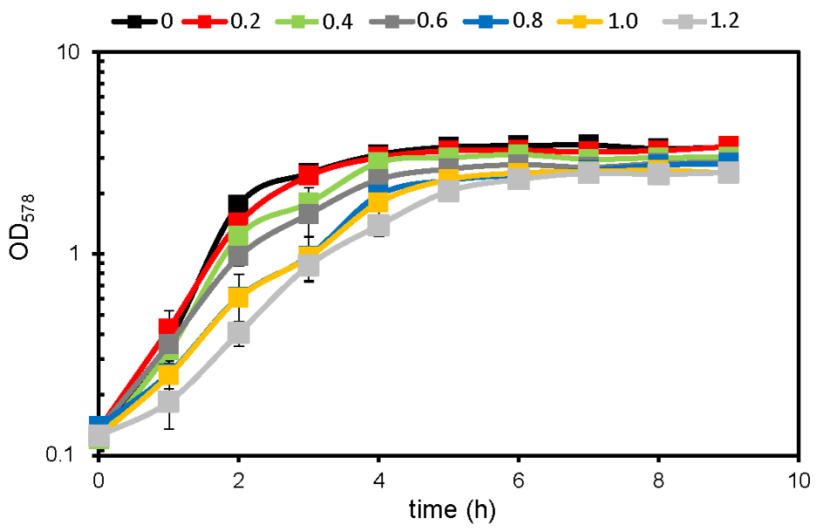


Figure S1. Growth of *Bacillus subtilis* in DSM in the presence of different concentrations of NaCl. Growth curves of *B. subtilis* JH642 cells grown in the presence of the indicated concentrations (in M) of NaCl in nutrient exhaustion medium (DSM). Lag-phase is extended in high salinity-challenged cells (above 0.6 M) but growth rate is only mildly impaired (less than 2-fold at 1.2 M NaCl).

Widderich *et al.*, Figure S2

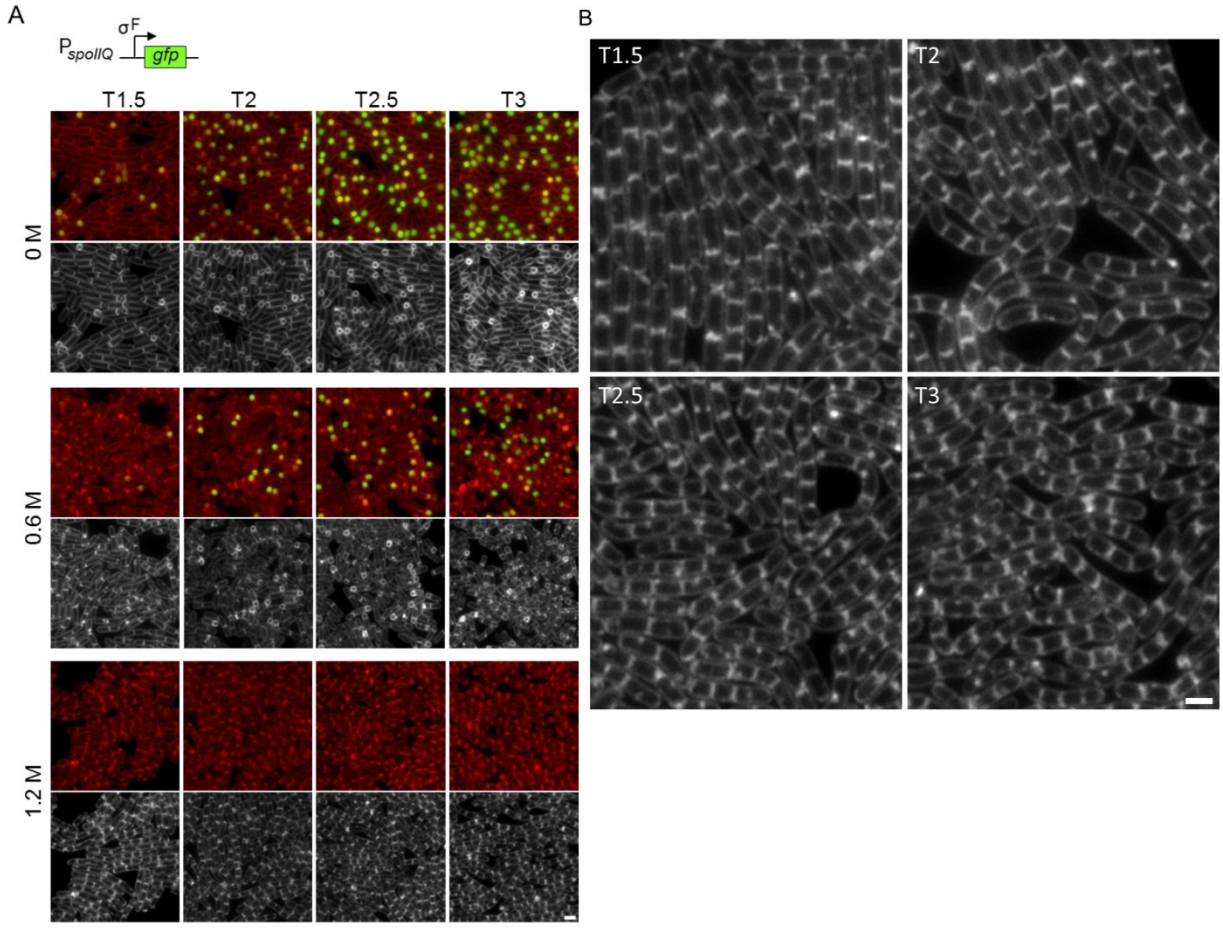


Figure S2. High salinity inhibits sporulation prior to asymmetric cell division. (A) Representative images of cells (strain BDR1048; Table S1) harboring a transcriptional reporter (P_{spoIIQ}-*gfp*) for the forespore-specific transcription factor σ^F induced to sporulate in the indicated concentrations of NaCl. Membranes (false-colored in red) were stained with TMA-DPH and σ^F-dependent expression of the reporter is shown in green. Membranes are also shown in black/white. The time after the initiation of sporulation is indicated. (B) Larger images of the same strain sporulated in the presence of 1.2 M NaCl to highlight the absence of polar septa. For comparison see Figure S9. Scale bar is 2 μm.

Widderich *et al.*, Figure S3

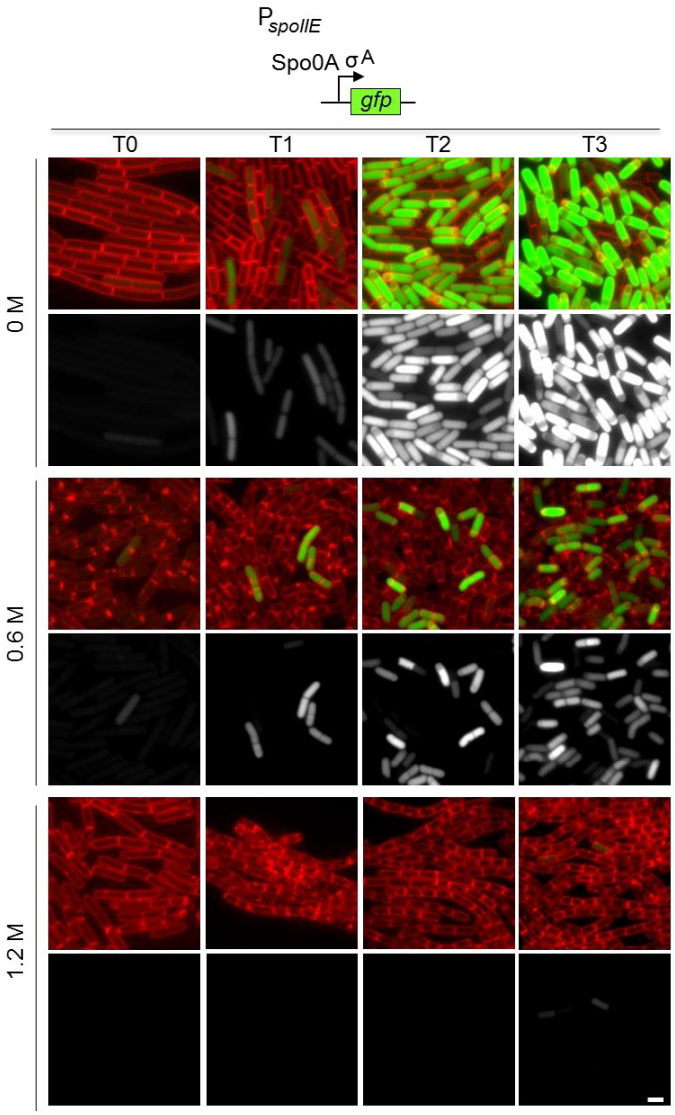


Figure S3. Spo0A activity is impaired in high salt. Representative images of cells (BDR2128) harboring a transcriptional reporter (P_{spoIIE}-*gfp*) for Spo0A activity induced to sporulate in the indicated concentrations of NaCl. For each condition and time point (in hours), GFP fluorescence is shown in black/white and above a merged image of membranes (red) and GFP (green). In the presence of 1.2 M NaCl no cells displayed Spo0A activity, whereas in 0.6 M NaCl a subset of cells displayed Spo0A activity. After polar division, Spo0A activity remains high in the mother cell compartment. All images were scaled identically. The relevant transcription factors that act on the *spoIIE* promoter are indicated above the images. Scale bar indicates 2 µm.

Widderich *et al.*, Figure S4

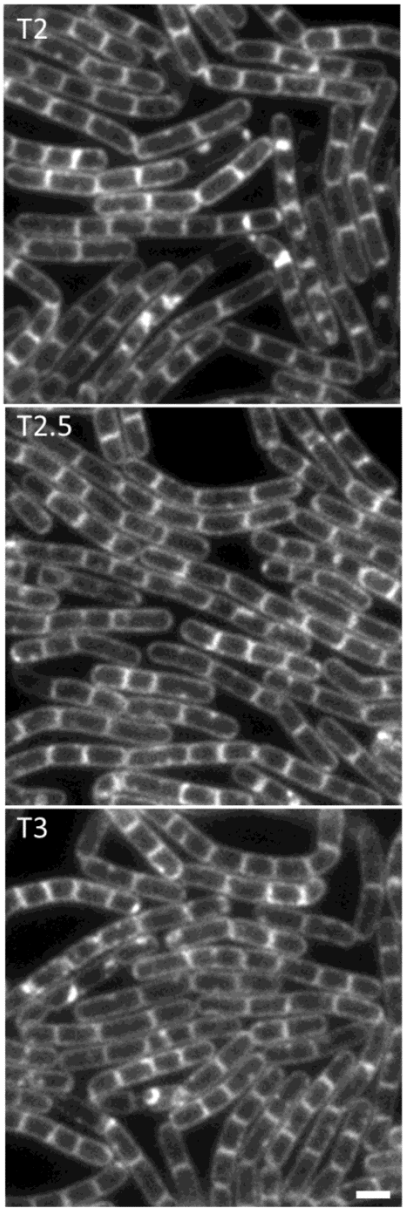


Figure S4. High salinity inhibits sporulation prior to asymmetric cell division. Representative images of cells (strain BDR2128; Table S1) harboring a transcriptional reporter (P_{spoIIE}-*gfp*) to monitor Spo0A-activity induced to sporulate in 1.2M NaCl. The hours after the initiation of sporulation are indicated. Membranes were stained with TMA-DPH and no asymmetric septa were observed at any of the indicated time points. Scale bar indicates 2 µm.

Widderich *et al.*, Figure S5

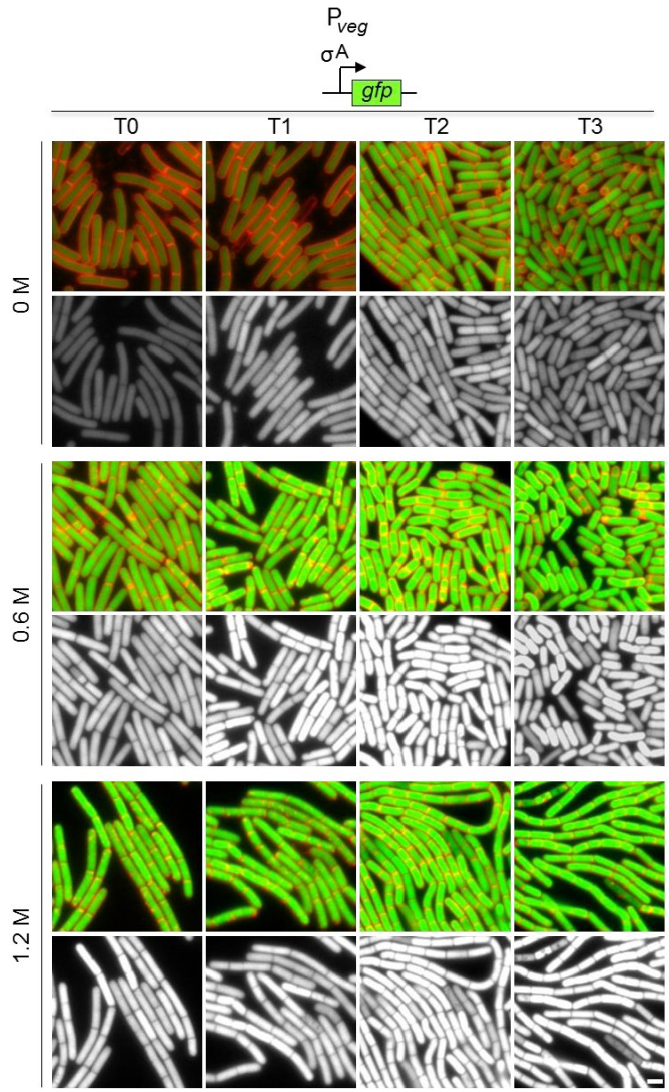


Figure S5. Control for GFP fluorescence and σ^A-dependent gene expression during sporulation in high salt. Representative images of cells (BDR2789) harboring a reporter (P_{veg}-*gfp*) for σ^A-dependent transcription induced to sporulate in the indicated concentrations of NaCl. For each condition and time point (in hours), GFP fluorescence is shown in black/white below a merged image of membranes (red) and GFP (green). GFP produced under σ^A control was easily detectable in 0.6M and 1.2M NaCl indicating that GFP fluorescence is not inhibited by high salt. Interestingly, σ^A-directed gene expression was higher in the presence of salt than in its absence, perhaps due to the absence of sigma factor competition. Scale bar indicates 2 µm.

Widderich *et al.*, Figure S6

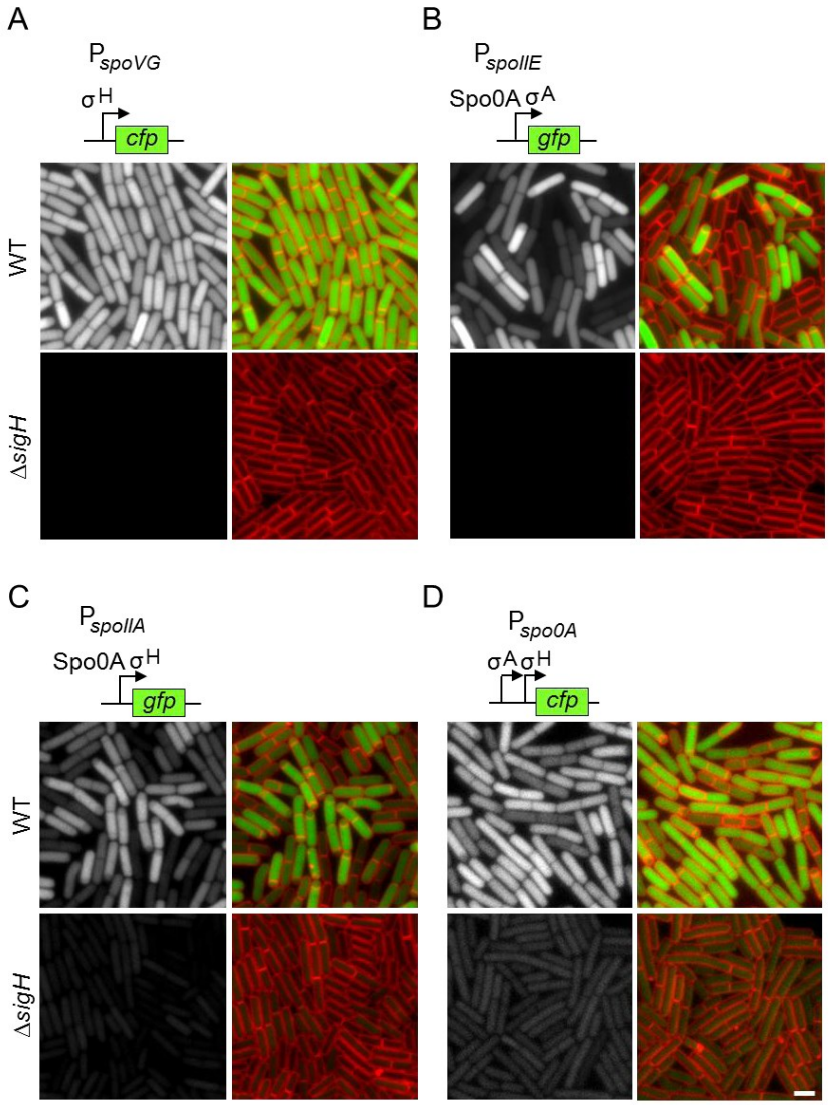


Figure S6. σ^H dependence of the transcriptional reporters used in this study. Representative images of sporulating cells at hour 1.5 harboring the indicated fluorescent reporters in wild-type and in cells lacking the gene (*sigH*) encoding σ^H. (A) The P_{spoVG}-*cfp* reporter in WT (top, BDR3064) and the *sigH* null (bottom, BDR3065), (B) the P_{spoIIE}-*gfp* reporter in WT (top, BDR2128) and the *sigH* null (bottom, BDR3074), (C) the P_{spoIIA}-*cfp* reporter in WT (top, BDR1799) and the *sigH* null (bottom, BDR3056) and (D) the P_{spo0A}-*cfp* reporter in WT (top, BDR3080) and the *sigH* null (bottom, BDR3083). Images are GFP/CFP-dependent fluorescence (black/white) and overlays of GFP/CFP (green) with membrane (red). Scale bar indicates 2 µm.

Widderich *et al.*, Figure S7

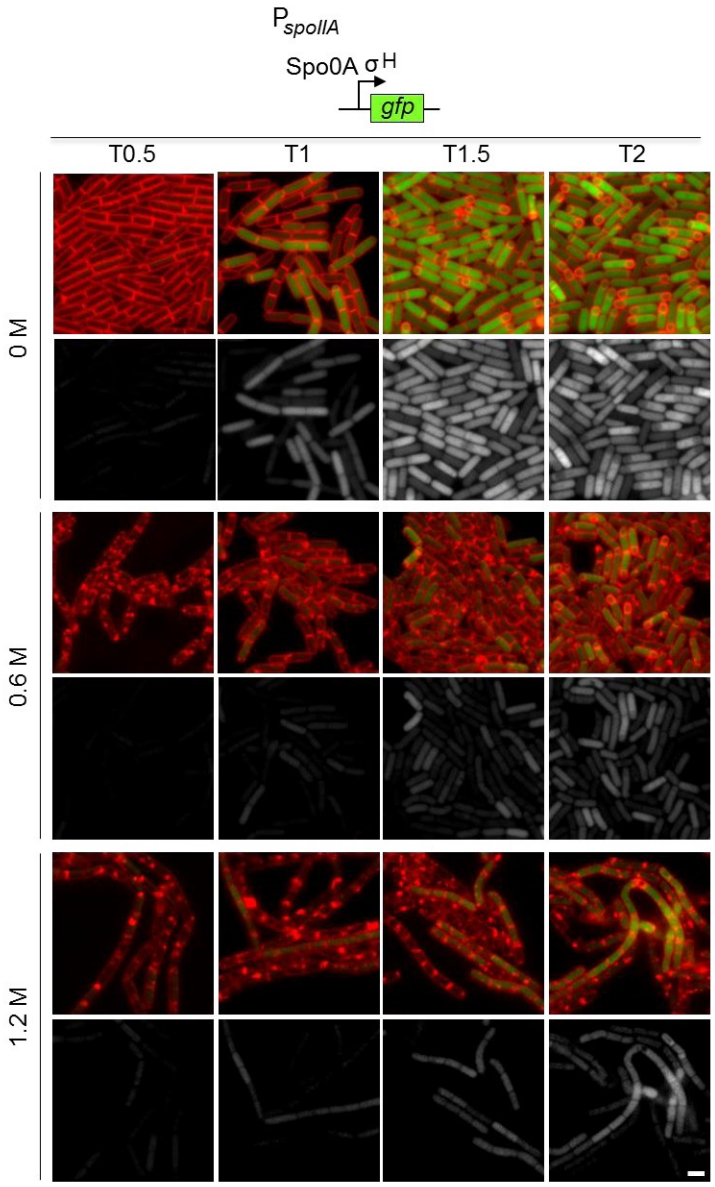


Figure S7. Transcription from the P_{spoIIA} promoter is reduced in the presence of high-salt. Representative images of cells (BDR1799) harboring a transcriptional reporter (P_{spoIIA}-*gfp*) for Spo0A- and σ^H-dependent activity induced to sporulate in the indicated concentrations of NaCl. For each condition and time point (in hours), GFP fluorescence is shown in black/white below a merged image of membranes (red) and GFP (green). Transcription from this promoter is reduced in 0.6 M NaCl and 1.2M NaCl. All images were scaled identically. The relevant transcription factors that act on the *spoIIA* promoter are indicated at the top. Scale bar indicates 2 µm.

Widderich *et al.*, Figure S8

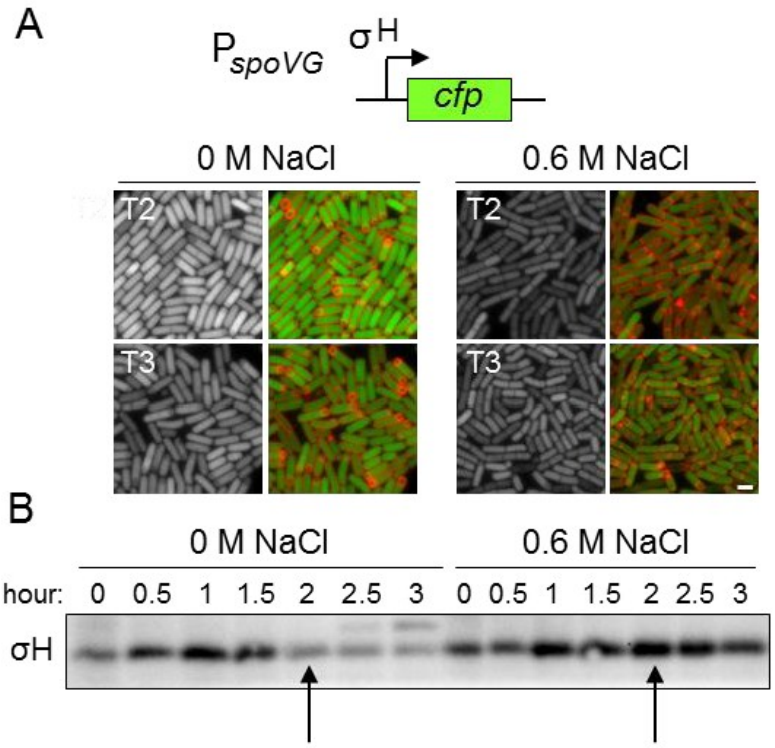


Figure S8. Comparison of σ^H activity and σ^H protein level in the presence and absence of 0.6 M NaCl. (A) Representative images of cells (strain BDR3064) harboring a transcriptional reporter (P_{spoVG}-*cfp*) for σ^H-dependent activity induced to sporulate in the indicated concentrations of NaCl and imaged at hours 2 (T2) and 3 (T3) after the onset of sporulation. All images were scaled identically. Scale bar indicates 2 μm. (B) Immunoblot analysis assessing σ^H levels in wild-type (strain PY79) induced to sporulate in the presence of the indicated concentrations of NaCl. Black arrows point to σ^H levels at the relevant time point for comparison to the images in (A). Although σ^H levels are higher in the 0.6 M NaCl condition, transcription from the σ^H-dependent promoter is lower.

Widderich *et al.*, Figure S9

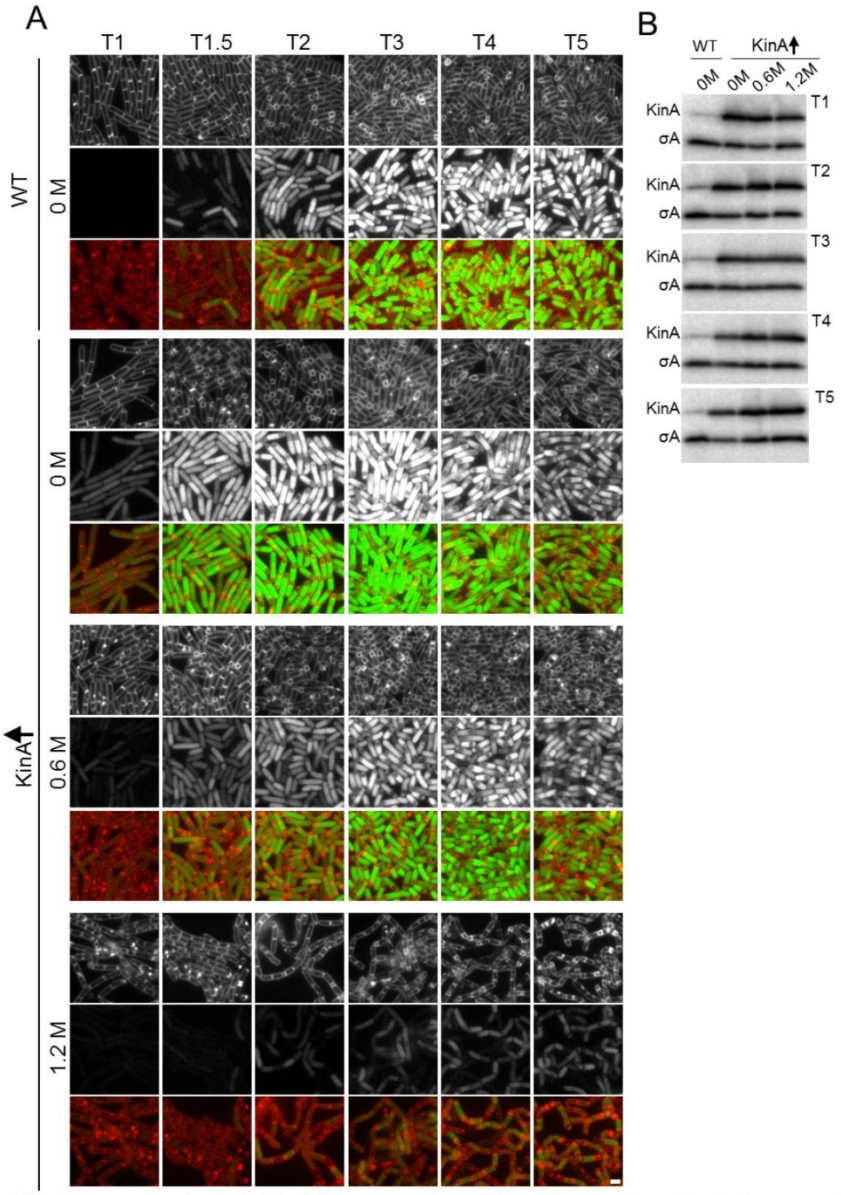


Figure S9. Overexpression of *kinA* bypasses the early salt-sensitive block in sporulation. (A) Representative images of cells harboring the Spo0A activity reporter (P_{spoIIE}-*gfp*) induced to sporulate in the presence of the indicated concentrations of NaCl. Images are from the wild-type (strain BDR2128) and a strain (BDR3087) harboring an IPTG-inducible allele of *kinA*. The time (in hours) after the initiation of sporulation is indicated. Membranes and GFP are shown in black/white with a merged image below. Scale bar indicates 2 μm. (B) Immunoblot analysis assessing the levels of KinA in the same strains in the indicated concentrations of NaCl. σ^A serves as a loading control.

Widderich *et al.*, Figure S10

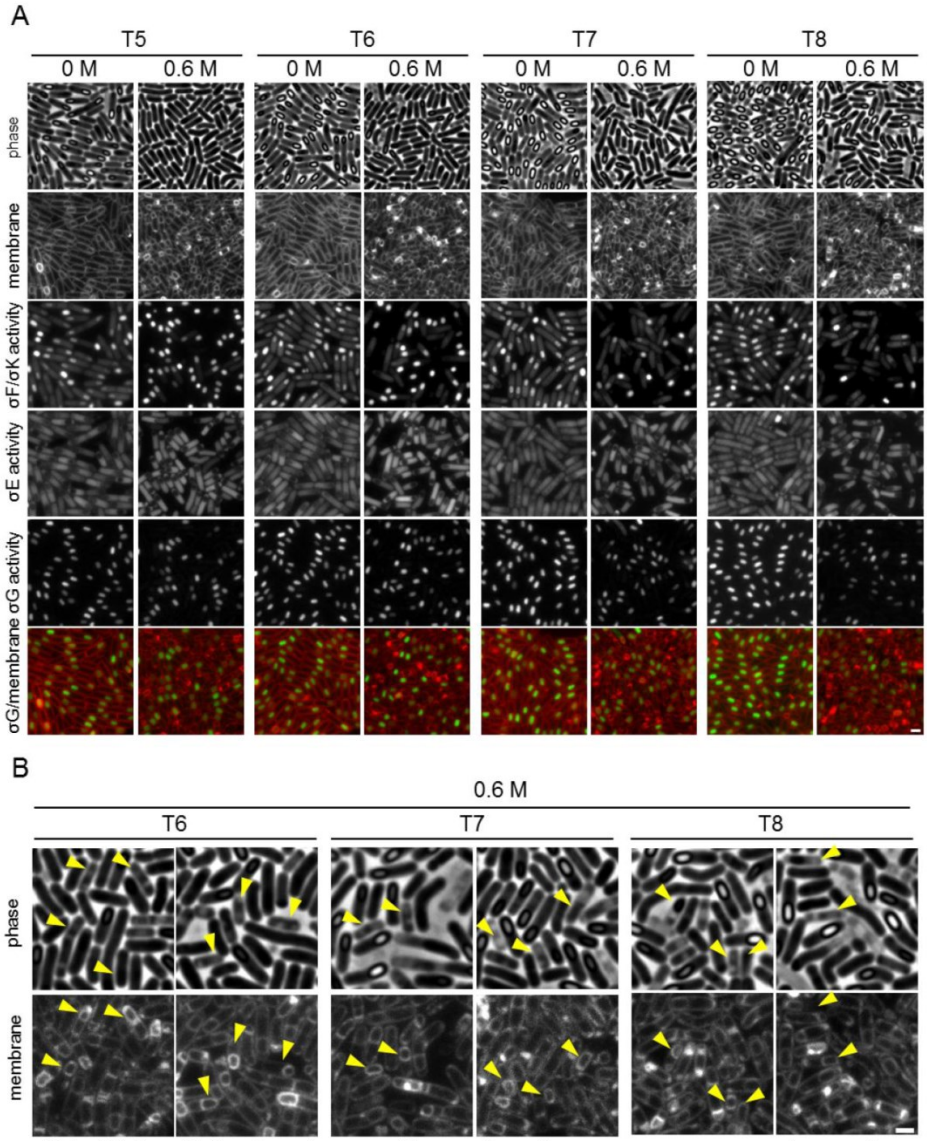


Figure S10. High salinity impairs engulfment, late forespore gene expression, reduces the production of phase bright spores and causes lysis. (A) Representative images of the *kinA* overexpression strain (BCR1274) harboring transcriptional reporters for all four stage- and compartment-specific sigma factors induced to sporulate in the presence of the indicated concentrations of NaCl (see main text for strain details). The time (in hours) after the initiation of sporulation is indicated. Phase panels highlight the reduction in phase bright spores and the increased lysis in cells sporulated in 0.6 M NaCl. (B) Representative images of the *kinA* overexpression strain highlighting the lysis of sporulating cells at the indicated time points. Yellow arrowheads point to sporulating cells that have lysed or are on the verge of lysing. Scale bar indicates 2 µm.

Widderich *et al.*, Figure S11

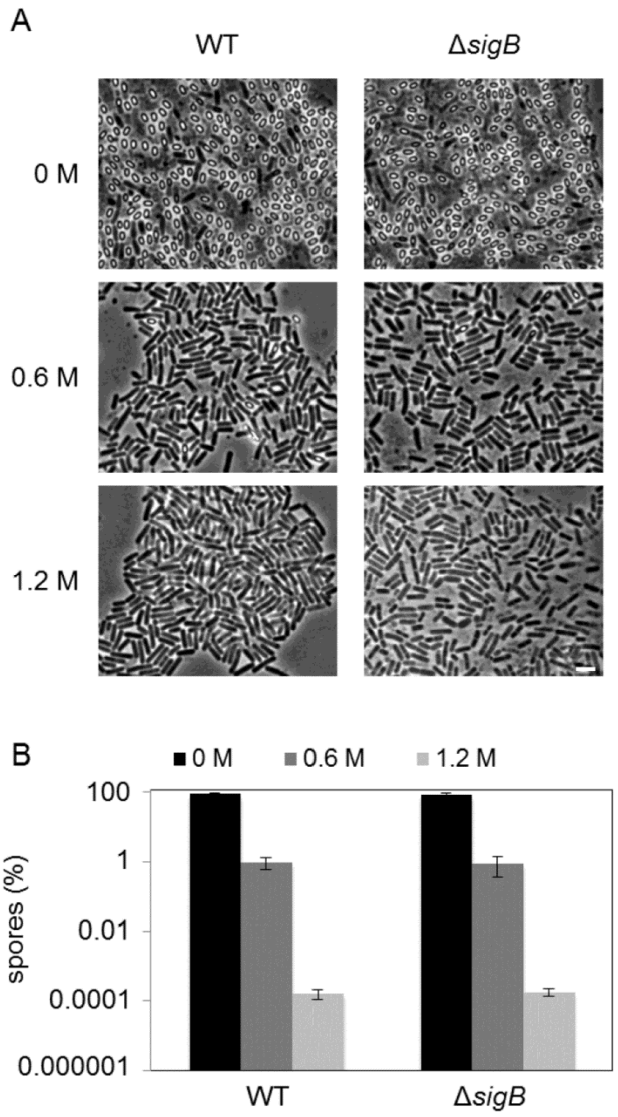


Figure S11. Sporulation at high salinity in a *sigB* null strain is similar to wild type. (A) Representative phase-contrast images of wild-type (WT; JH642) and a *sigB* null mutant (FSB5) sporulated by nutrient exhaustion in the presence of the indicated concentrations of NaCl. Scale bar indicates 4 µm. (B) The bar graph shows the negative effects of high salinity on sporulation.

Widderich *et al.*, Figure S12

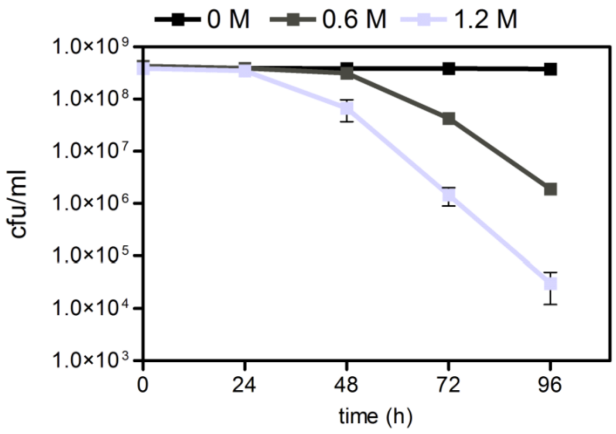


Figure S12. Survival of *Bacillus subtilis* cells grown in the presence of different concentrations of NaCl. The colony forming units per ml (cfu/ml) of *B. subtilis* JH642 cells incubated for the indicated times (in hours) in nutrient exhaustion medium (DSM) in presence of 0 M, 0.6 M or 1.2 M NaCl. In the presence of 0 M NaCl, the cells form dormant spores and the colony forming units remain constant over 96 hours. Cells grown in medium with high salt lose viability after 48-72 hours.

Widderich *et al.*, Table S1

Table S1. Strains used in this study

strain	genotype	source
JH642	*trpC2 pheA1*	(Dean *et al.*, 1977)
168	*trpC2*	(Burkholder and Giles, 1947)
3610	undomesticated wild type (biofilm strain)	(Branda *et al.*, 2001)
DK1042	comI^{Q12L}	(Konkol *et al.*, 2013)
PY79	Prototrophic wild-type strain	(Youngman *et al.*, 1983)
B. mojavensis	wild-type strain	(Earl *et al.*, 2012)
FSB5	*sigBD2::spc*	(Spiegelhalter and Bremer, 1998)
BDR1753	*sigH::kan, trpC2, pheA1*	D. Rudner
BDR1048	*amyE::PspoIIQ-gfp (spec)*	D. Rudner
BDR2128	*amyE::PspoIIE-gfp (spec)*	(Fujita and Losick, 2003)
BDR2331	*kinaAΩPhyper-spank-kinA(spec)*	(Fujita and Losick, 2005)
BDR2789	*sacA::Pveg-gfp (phleo)*	D. Rudner
BDR485	*rpoCΩpETΔrpoC (rpoC-his6) (neo)*	(Fujita and Sadaie, 1998)
BCR1071	*yycR::PsspB-rbsopt-cfp (phleo), amyE::PspoIID-mCherry(spec), pelB::PspoIIQ-yfp (kan), lacA::PgerE-yfp (tet)*	(Meeske *et al.*, *in press*)
BCR1274	*yycR::PsspB-rbsopt-cfp (phleo), amyE::PspoIID-mCherry(Bsub OPT)(spec), pelB::PspoIIQ-yfp (kan), lacA::PgerE-yfp (tet), kinAΩPhyperspank-kinA (cat)*	this study
BDR1799	*amyE::PspoIIA-gfp (cat)*	D. Rudner
BDR3056	*sigHΔ, amyE::PspoIIA-gfp cat*	this study
BDR3064	*amyE::PspoVG-cfp (spec)*	this study
BDR3065	*sigH::kan, amyE::PspoVG-cfp (spec)*	this study
BDR3074	*sigH::kan, amyE::PspoIIE-gfp (spec)*	this study
BDR3080	*amyE::Pspo0A-cfp (spec)*	this study
BDR3083	*sigH::kan, amyE::Pspo0A-cfp (spec)*	this study
BDR3087	*kinaAΩPhyper-spank-kinA(cat), amyE::PspoIIE-gfp (spec)*	this study
BDR3090	*amyE::PsigH-cfp (spec)*	this study
BDR3095	*sigH::kan, amyE::PsigH-cfp (spec)*	this study
NWB6	JH642-derived suppressor allowing increased sporulation at high salinity	this study
NWB7	JH642-derived suppressor allowing increased sporulation at high salinity	this study
NWB11	JH642-derived suppressor allowing increased sporulation at high salinity	this study
NWB13	JH642-derived suppressor allowing increased sporulation at high salinity	this study
NWB16	JH642-derived suppressor allowing increased sporulation at high salinity	this study
NWB17	JH642-derived suppressor allowing increased sporulation at high salinity	this study
NWB19	JH642-derived suppressor allowing increased sporulation at high salinity	this study
NWB21	JH642-derived suppressor allowing increased sporulation at high salinity	this study
NWB22	JH642-derived suppressor allowing increased sporulation at high salinity	this study
NWB24	JH642-derived suppressor allowing increased sporulation at high salinity	this study
NWB25	JH642-derived suppressor allowing increased sporulation at high salinity	this study

Widderich *et al.*, Table S2

Table S2. Plasmids used in this study

plasmids	description	source
pKM161	*amyE::*P_{spoIIQ}*-cfp XhoI (spec) (amp)*	D. Rudner
pNW1	*amyE::*P_{sigH}*-cfp (spec) (amp)*	this study
pNW2	*amyE::*P_{spoVG}*-cfp (spec) (amp)*	this study
pNW4	*amyE::*P_{spo0A}*-cfp (spec) (amp)*	this study

Widderich *et al.*, Table S3

Table S3. Oligonucleotide primers used in this study

primer	sequence
oDR1200	ccatagtagttcctccttatgtaagctttttccctatataaaagcattagtgtatc
oDR1202	tcgccattcgccagggctgcaggaattcgaaataataagaaaagtgattctggg
oDR1203	ttattgccaaaagagctatccacgc
oDR1204	aacaacggcgtttaagccgttgttt
oDR1205	attttcggacaggggcattgcgga
oDR1206	actgtcaatatagcataaattccta
oDR1207	tattttgatccctcttcacttctcag
oDR1208	taatgaccacttaataagctcatgt
oDR75	ataacatgtattcacgaacg
oDR1217	cgcgaattcttatggaaaagaaaaagcaagctgactgccgg
oDR1218	cgcaagcttaattttctccacgtttcttcctcccc
oDR1219	gcgaagcttatggtttcaaaaggcgaagaactg
oDR1220	cgcggatcctagcatgcatgctagcat
oDR1221	cgcgaattcagagaggtagaaacgattgaaaggcg
oDR1222	cgcaagcttctgtagattcactccgatcccc
oML79	ctcttgccagtcacgttacg
oDR77	ggcagacatggcctgcccgg

Supplementary methods

Strain construction

All *B. subtilis* strains were constructed by standard transformation with genomic DNA or linearized plasmids followed by selection on LB agar plates supplemented with appropriate antibiotics.

Plasmid construction

pNW1 [amyE::P_{sigH}-*cfp* (spec)] was obtained by amplification of P_{sigH} from genomic DNA of strain PY79 using primers oDR1221 and oDR1222 (Table S3) and by amplification of *cfp* from plasmid pKM161 (Table S2) using oligonucleotides oDR1219 and oDR1220 (Table S3). The P_{sigH} product was restricted with EcoRI and HindIII, the *cfp* product with HindIII and BamHI and the pKM161 plasmid with EcoRI and BamHI. Restricted products and backbone of plasmid pKM161 were purified from agarose gels and then ligated in a three-way ligation using T4 DNA Ligase. The resulting plasmid was transformed into *E. coli* DH5α and plated on LB agar plates containing ampicillin (100 µg ml^{-1}). DNA sequencing with primers oML79 & oDR77 (Table S3) was performed to confirm the structure of the desired P_{sigH}-*cfp* reporter plasmid (Table S2).

pNW2 [amyE::P_{spoVG}-*cfp* (spec)] was obtained by isothermal assembly (Gibson, 2011). The isothermal assembly reaction contained two PCR products: (i) P_{spoVG} (amplified from genomic DNA of strain PY79 using primers oDR1200 and oDR1202; Table S3), and (ii) plasmid pKM161 that was restricted with EcoRI-HF and HindIII-HF. The resulting plasmid was transformed into *E. coli* DH5α and plated on LB agar plates containing ampicillin. Sequencing with primer oDR75 (Table S3) was performed to confirm the structure of the desired P_{spoVG}-*cfp* reporter construct (Table S2).

pNW4 [amyE::P_{spo0A}-*cfp* (spec)] was obtained by amplification of P_{spo0A} from chromosomal DNA of strain PY79 using primers oDR1217 and oDR1218 (Table S3) and by amplification of *cfp* from plasmid pKM161 (Table S2) using oligonucleotides oDR1219 and oDR1220 (Table S3). The resulting P_{spo0A} product was restricted with EcoRI and HindIII, the *cfp* product with HindIII and BamHI and plasmid pKM161 with EcoRI and BamHI. Restricted products and backbone of pKM161 were purified from agarose gels and combined in a three-way ligation using T4 DNA Ligase. The resulting plasmid was transformed into *E. coli* DH5α and plated on LB agar plates containing ampicillin. Sequencing with primers oML79 & oDR77 (Table S3) was performed to confirm the structure of the desired P_{spo0A}-*cfp* reporter construct (Table S2).

References

Branda, S.S., Gonzalez-Pastor, J.E., Ben-Yehuda, S., Losick, R., and Kolter, R. (2001) Fruiting body formation by *Bacillus subtilis*. *Proc Natl Acad Sci U S A* **98**: 11621-11626.

Burkholder, P.R., and Giles, N.H., Jr. (1947) Induced biochemical mutations in *Bacillus subtilis*. *Am J Bot* **34**: 345-348.

Dean, D.R., Hoch, J.A., and Aronson, A.I. (1977) Alteration of the *Bacillus subtilis* glutamine synthetase results in overproduction of the enzyme. *J Bacteriol* **131**: 981-987.

Earl, A.M., Eppinger, M., Fricke, W.F., Rosovitz, M.J., Rasko, D.A., Daugherty, S., Losick, R., Kolter, R., and Ravel, J. (2012) Whole-genome sequences of *Bacillus subtilis* and close relatives. *J Bacteriol* **194**: 2378-2379.

Fujita, M., and Losick, R. (2003) The master regulator for entry into sporulation in *Bacillus subtilis* becomes a cell-specific transcription factor after asymmetric division. *Genes & Development* **17**: 1166-1174.

Fujita, M., and Losick, R. (2005) Evidence that entry into sporulation in *Bacillus subtilis* is governed by a gradual increase in the level and activity of the master regulator Spo0A. *Genes & Development* **19**: 2236-2244.

Fujita, M., and Sadaie, Y. (1998) Rapid isolation of RNA polymerase from sporulating cells of *Bacillus subtilis*. *Gene* **221**: 185-190.

Gibson, D.G. (2011) Enzymatic assembly of overlapping DNA fragments. *Methods Enzymol* **498**: 349-361.

Konkol, M.A., Blair, K.M., and Kearns, D.B. (2013) Plasmid-encoded ComI inhibits competence in the ancestral 3610 strain of *Bacillus subtilis*. *J Bacteriol* **195**: 4085-4093.

Meeske, A.J., Rodrigues, C.D.A., Brady, J., Lim, H.C., Bernhardt, T.G., and Rudner, D.Z. (2015) High throughput genetic screens identify a large and diverse collection of new sporulation genes in *Bacillus subtilis*. PLoS Biology, *in press*.

Spiegelhalter, F., and Bremer, E. (1998) Osmoregulation of the *opuE* proline transport gene from *Bacillus subtilis*: contributions of the sigma A- and sigma B-dependent stress-responsive promoters. *Molecular Microbiology* **29**: 285-296.

Youngman, P.J., Perkins, J.B., and Losick, R. (1983) Genetic transposition and insertional mutagenesis in *Bacillus subtilis* with *Streptococcus faecalis* transposon Tn917. *Proc Natl Acad Sci U S A* **80**: 2305-2309.

Danksagung

An erster Stelle möchte ich Herrn Prof. Dr. Erhard Bremer für die Überlassung des interessanten Themas und die Begutachtung der vorliegenden Arbeit bedanken.

Herrn Prof. Dr. Andreas Brune danke Ich für die Bereitschaft, als Zweitgutachter für die vorliegende Arbeit zu fungieren sowie mich für die dreijährige Begleitung meiner Arbeit im Rahmen meines IMPRS-Thesis-Komitees bedanken. In diesem Zusammenhang gilt mein Dank ebenfalls den weiteren Mitgliedern des Komitees, Dr. Chris van der Does und Dr. Tobias Erb. An dieser Stelle sei auch der IMPRS für die dreijährige Förderung und Finanzierung gedankt. Gleiches gilt auch für Herrn Prof. Rolf Thauer, für dessen Unterstützung ich Danke sagen möchte. Zudem möchte ich mich bei den weiteren Mitgliedern meiner Prüfungskommission, Prof. Dr. Hans-Ulrich Mösch und Prof. Dr. Uwe Meier, bedanken.

Ganz großen Dank möchte ich Dr. Sander Smits von der Universität Düsseldorf für die außergewöhnlich fruchtbare und tolle Kollaboration aussprechen. Ohne deine Mühen, unsere unzähligen Mails und deine stetige Unterstützung wäre diese Arbeit niemals so erfolgreich gewesen. Vielen Dank! In diesem Zusammenhang möchte ich mich auch bei seinen Mitarbeitern bedanken, deren großes Engagement die Auflösung der Kristallstrukturen erst ermöglichte.
In diesem Zusammenhang möchte ich auch Herrn Prof. Dr. Wolfgang Buckel, Prof. Dr. Johann Heider, Prof. Dr. Jeroen Dickschat, Dr. Martin Könneke und allen weiteren Kollaborationspartnern meinen Dank aussprechen.

Hervorheben möchte ich dabei David Rudner von der Harvard Medical School, der mir ermöglicht hat, in seinem Labor an der HMS und unter seiner kompetenten Leitung das Sporulationsprojekt unter den besten Gegenebenheiten und in einer unglaublichen Umgebung zu bearbeiten. Dieser Aufenthalt war eine ganz besondere, inspirierende und unvergessliche Erfahrung. Dafür danke ich dir! Dabei möchte ich Chris Rodrigues ein ganz großes Dankeschön aussprechen, der mich während meines Aufenthalts in Boston nicht nur im Labor stets unterstützt hat und mir mit Rat und Tat zur Seite stand, sondern auch darüber hinaus. Es war mir eine große Freude mit dir zusammenzuarbeiten! Ebenfalls möchte ich mich bei allen Mitgliedern aus Davids Arbeitsgruppe für die tolle Atmosphäre und die wunderschöne Zeit bedanken. In diesem Zusammenhang möchte ich auch der EMBO Gesellschaft großen Dank aussprechen, die mir durch die Vergabe eines „Short-term fellowship“ den Aufenthalt in Boston erst ermöglichte.

Ein besonderer Dank gebührt Dr. Tamara Hoffman für ihre stetigen Hilfestellungen, ihre unglaubliche Geduld, sowie ihren Rat und ihre Tipps bei einer Vielzahl wissenschaftlicher und anderer Fragestellungen während all der Zeit. Dabei möchte ich ihr vor allem für ihre Ehrlichkeit und konstruktive Kritik danken. Hervorheben möchte ich auch Jochen Sohn für seine Unterstützung und seine Hilfe in den letzten drei Jahren. Es war mir eine Freude mit dir zusammenzuarbeiten und ich werde es definitiv sehr vermissen. Danke JJ! Darüber hinaus möchte ich auch allen weiteren Labormitarbeitern für den tollen Laboralltag, die Kaffeeklätschen und die schöne Zeit bedanken. Zudem möchte ich Jochen und Marco Hornung für ihre Hilfestellungen in zahlreichen technischen Problemen und die vertiefte Einführung in die Flüssigkeitschromatographie danken!

Natürlich möchte ich mich auch bei der gesamten Mikrobiologie (AG Bremer, AG Heider und AG Buckel) für die familiäre und harmonische Atmosphäre während der ganzen Zeit bedanken sowie für die netten „Grill-Nachmittage“, die in dieser Zeit stattgefunden haben.

Nicht zuletzt gilt mein besonderer Dank meiner Familie, meinen Eltern Hiltrud und Cay Widderich, sowie meiner Schwester Ilka, die mich in allen Lebenslagen bedingungslos unterstützt und mir Rückhalt gegeben haben, immer für mich da waren und damit vieles erst ermöglich haben!

Ganz besonders danken möchte ich meiner Frau Lisa, die mich während dieser Zeit und darüber hinaus stets unterstützt und mir uneingeschränktes Verständnis entgegengebracht hat sowie mir stets zur Seite stand.

Vielen Dank!

Erklärung

Hiermit versichere ich, dass ich die vorliegende Dissertation mit dem Titel:

„**Phylogenetic, biochemical and structural assessment of key enzymes in ectoine and hydroxyectoine biosynthesis**“

selbständig verfasst und keine anderen als die im Text angegebenen Hilfsmittel verwendet habe. Sämtliche Textstellen, die im Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind mit einer Quellenangabe kenntlich gemacht.

Die Dissertation wurde in der jetzigen oder ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Nils Widderich	Ort, Datum